

Thu Jun 26 11:58:18 2003

us-09-980-370-8.rag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 26, 2003, 00:44:59 ; Search time 7.4117 seconds
(without alignments)
377.543 Million cell updates/sec

Title: US-09-980-370-8

Sequence: 1 ALNPDVLIKSDVIGATKRYK 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/emb1/AA1980.DAT:
2: /SID2/gcgdata/geneseq/emb1/AA1981.DAT:
3: /SID2/gcgdata/geneseq/emb1/AA1982.DAT:
4: /SID2/gcgdata/geneseq/emb1/AA1983.DAT:
5: /SID2/gcgdata/geneseq/emb1/AA1984.DAT:
6: /SID2/gcgdata/geneseq/emb1/AA1985.DAT:
7: /SID2/gcgdata/geneseq/emb1/AA1986.DAT:
8: /SID2/gcgdata/geneseq/emb1/AA1987.DAT:
9: /SID2/gcgdata/geneseq/emb1/AA1988.DAT:
10: /SID2/gcgdata/geneseq/emb1/AA1989.DAT:
11: /SID2/gcgdata/geneseq/emb1/AA1990.DAT:
12: /SID2/gcgdata/geneseq/emb1/AA1991.DAT:
13: /SID2/gcgdata/geneseq/emb1/AA1992.DAT:
14: /SID2/gcgdata/geneseq/emb1/AA1993.DAT:
15: /SID2/gcgdata/geneseq/emb1/AA1994.DAT:
16: /SID2/gcgdata/geneseq/emb1/AA1995.DAT:
17: /SID2/gcgdata/geneseq/emb1/AA1996.DAT:
18: /SID2/gcgdata/geneseq/emb1/AA1997.DAT:
19: /SID2/gcgdata/geneseq/emb1/AA1998.DAT:
20: /SID2/gcgdata/geneseq/emb1/AA1999.DAT:
21: /SID2/gcgdata/geneseq/emb1/AA2000.DAT:
22: /SID2/gcgdata/geneseq/emb1/AA2001.DAT:
23: /SID2/gcgdata/geneseq/emb1/AA2002.DAT:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	21	AA849218	Peptide #3. Unide
2	108	100.0	134	AA849217	Peptide used in th
3	108	100.0	439	AA86024	P. gingivalis haem
4	108	100.0	439	AA86024	Haemagglutinin pro
5	108	100.0	450	AA86021	P. gingivalis haem
6	108	100.0	450	AA86021	Haemagglutinin pro
7	108	100.0	456	AA86022	P. gingivalis haem
8	108	100.0	456	AA86022	Haemagglutinin pro
9	108	100.0	456	AA86023	P. gingivalis haem
10	108	100.0	456	AA86023	Haemagglutinin pro

11	108	100.0	1087	17	AA86028	P. gingivalis haem
12	108	100.0	1087	19	AA86028	Haemagglutinin pro
13	108	100.0	1358	17	AA86032	P. gingivalis haem
14	108	100.0	1358	19	AA86032	Haemagglutinin pro
15	108	100.0	1687	17	AA86033	P. gingivalis haem
16	108	100.0	1687	19	AA86033	Haemagglutinin pro
17	108	100.0	1704	16	AA870188	Arg-gingipain-2 pr
18	108	100.0	1704	16	AA870188	Arg-gingipain-2 pr
19	108	100.0	1704	21	AA870188	Arg-gingipain-2 am
20	108	100.0	1704	21	AA870188	Arg-gingipain-2 am
21	108	100.0	1706	22	AA870188	P. gingivalis high
22	108	100.0	1706	22	AA870188	P. gingivalis high
23	108	100.0	1732	17	AA86029	P. gingivalis porp
24	108	100.0	1732	18	AA86029	P. gingivalis porp
25	108	100.0	1732	19	AA86029	P. gingivalis porp
26	108	100.0	1732	19	AA86029	P. gingivalis porp
27	108	100.0	1732	19	AA86029	P. gingivalis porp
28	108	100.0	1732	19	AA86029	P. gingivalis porp
29	108	100.0	1732	19	AA86029	P. gingivalis porp
30	108	100.0	1732	19	AA86029	P. gingivalis porp
31	108	100.0	1732	19	AA86029	P. gingivalis porp
32	108	100.0	1732	19	AA86029	P. gingivalis porp
33	108	100.0	1732	19	AA86029	P. gingivalis porp
34	108	100.0	1732	19	AA86029	P. gingivalis porp
35	108	100.0	1732	19	AA86029	P. gingivalis porp
36	108	100.0	1732	19	AA86029	P. gingivalis porp
37	108	100.0	1732	19	AA86029	P. gingivalis porp
38	108	100.0	1732	19	AA86029	P. gingivalis porp
39	108	100.0	1732	19	AA86029	P. gingivalis porp
40	108	100.0	1732	19	AA86029	P. gingivalis porp
41	108	100.0	1732	19	AA86029	P. gingivalis porp
42	108	100.0	1732	19	AA86029	P. gingivalis porp
43	108	100.0	1732	19	AA86029	P. gingivalis porp
44	108	100.0	1732	19	AA86029	P. gingivalis porp
45	108	100.0	1732	19	AA86029	P. gingivalis porp

ALIGNMENTS

RESULT 1	AA849218	standard; peptide; 21 AA.
ID	AA849218	standard; peptide; 21 AA.
AC	AA849218	
DT	13-MAR-2001	(First entry)
DE	Peptide #3.	
KM	HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;	
OS	Infection.	
PS	Unidentified.	
PN	WO200072875-A1.	
PD	07-DEC-2000.	
PE	26-MAY-2000; 2000WO-A000599.	
PR	28-MAY-1999; 99AU-0000652.	
PA	(UNSY) UNIV SYDNEY.	
PI	Collyer CA, Hunter N, De Carlo AA;	
DR	WPI: 2001-080424/09.	
PT	Treating microbial infection in environment containing porphyrin, by	
PS	administering a HA-2 antagonist	
Claim 11; Page 40; 102pp; English.		


```

XX Haemagglutinin protein; periodontal disease; vaccine; hga.
XX Porphyromonas gingivalis.
XX US5824791-A.
XX 20-OCT-1998.
XX 11-DEC-1995; 95US-0570311.
XX 11-DEC-1995; 95US-0570311.
XX 08-SEP-1988; 88US-0241640.
XX 25-JAN-1991; 91US-0647119.
XX 09-DEC-1994; 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.
XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
XX Tumwasorn S;
XX WPI: 1998-582627/49.
XX N-PSDB: AAV58879.
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX and/or protease poly(peptide(s))
XX Claim 1: Column 139-144; 101pp; English.
XX This sequence is encoded by a Porphyromonas gingivalis gene of the
XX invention. This sequence represents the hga haemagglutinin protein. The
XX polypeptides are used to produce antibodies to a gingivitis associated with
XX periodontal disease. The antibodies are also used in purification and
XX identification procedures. The genes and polypeptides are used as
XX vaccines against periodontal disease.
XX Sequence 439 AA:
XX
XX Query Match 100.0%; Score 108; DB 19; Length 439;
XX Best Local Similarity 100.0%; Pred. No. 2e-10;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ALNPDNYLSKDYGTATKVKY 21
XX 251 ALNPDNYLSKDYGTATKVKI 271
XX
XX Db
XX
XX RESULT 5
XX AAR96021 standard; Protein: 450 AA.
XX AAR96021:
XX 04-SEP-1996 (first entry)
XX P. gingivalis haemagglutinin hga HAREPI product.
XX Haemagglutinin; hga; periodontal disease; vaccine; antibody;
XX HAREPI.
XX Porphyromonas gingivalis strain 381.
XX WO9617936-A2.
XX 13-JUN-1996.
XX 11-DEC-1995; 95MO-US16108.
XX 09-DEC-1994; 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.

```

```

XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
XX Tumwasorn S;
XX WPI: 1996-287181/29.
XX N-PSDB: AAT30645.
XX Porphyromonas gingivalis genes and proteins - used in the detection
XX and vaccination against periodontal disease
XX Claim 4: Page 103-104; 153pp; English.
XX HAREPI (AAR96021) is the product of the HAREPI repeat unit (AAT30645)
XX of the hga gene (AAT30654) of P. gingivalis 318. It forms part
XX of haemagglutinin hga (see also AAR96030). HAREPI and other
XX hga repeat unit products (see also AAR96022-24) can be obtd. from
XX transformed host cells and used as vaccines to protect humans or
XX animals against periodontal disease. Expression in Salmonella
XX cells allows produ. of live vaccine. HAREPI-4 can also be used
XX to detect the presence of anti-P. gingivalis antibodies and to
XX raise monoclonal antibodies for diagnostic appln.
XX Sequence 450 AA:
XX
XX Query Match 100.0%; Score 108; DB 17; Length 450;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-10;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ALNPDNYLSKDYGTATKVKY 21
XX 245 ALNPDNYLSKDYGTATKVKY 265
XX
XX Db
XX
XX RESULT 6
XX AAR96489 standard; Protein: 450 AA.
XX AAR96489:
XX 22-DEC-1998 (first entry)
XX Haemagglutinin protein hga, HAREPI.
XX Haemagglutinin protein; periodontal disease; vaccine; hga.
XX Porphyromonas gingivalis.
XX US5824791-A.
XX 20-OCT-1998.
XX 11-DEC-1995; 95US-0570311.
XX 11-DEC-1995; 95US-0570311.
XX 08-SEP-1988; 88US-0241640.
XX 25-JAN-1991; 91US-0647119.
XX 09-DEC-1994; 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.
XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
XX Tumwasorn S;
XX WPI: 1998-582627/49.
XX N-PSDB: AAV58879.
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX and/or protease poly(peptide(s))
XX Claim 1: Column 121-126; 101pp; English.
XX This sequence is encoded by a Porphyromonas gingivalis gene of the

```

CC Invention. This sequence represents the hga haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

XX Sequence 450 AA:

Query Match 100.0%; Score 108; DB 19; Length 450;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNLTISKDVGTATKVKY 21
 DB 245 ALNPDNLTISKDVGTATKVKY 265

RESULT 7
 AAR96022 standard; Protein: 456 AA.

XX AAR96022:

XX 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hga Harep2 product.

XX Haemagglutinin: hga; periodontal disease; vaccine; antibody;

XX Harep2.

XX Porphyromonas gingivalis strain 381.

XX WO9617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995; 95WO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Progulskie-Fox A;

XX Tumwasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB: AAT30646.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4; Page 107-108; 153pp; English.

XX Harep2 (AAR96022) is the product of the Harep2 repeat unit (AAT30646)

XX of the hga gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin hga (see also AAR96030). Harep2 and other hga

XX repeat unit products (see also AAR96021 and AAR96023-24) can be obtd.

XX from transformed host cells and used as vaccines to protect humans

XX or animals against periodontal disease. Expression in Salmonella

XX cells allows prodn. of live vaccine. Harep1-4 can also be used

XX to detect the presence of anti-P. gingivalis antibodies and to

XX raise monoclonal antibodies for diagnostic appln.

XX Sequence 456 AA:

Query Match 100.0%; Score 108; DB 17; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNLTISKDVGTATKVKY 21
 DB 251 ALNPDNLTISKDVGTATKVKY 271

RESULT 8
 AAR96023 standard; Protein: 456 AA.

XX AAR96023:

XX 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hga Harep3 product.

XX Haemagglutinin: hga; periodontal disease; vaccine; antibody;

XX Harep3.

XX Porphyromonas gingivalis strain 381.

XX WO9617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995; 95WO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Progulskie-Fox A;

XX Tumwasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB: AAT30647.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4; Page 110-112; 153pp; English.

XX Harep3 (AAR96023) is the product of the Harep3 repeat unit (AAT30647)

XX of the hga gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin hga (see also AAR96030). Harep3 and other hga

XX repeat unit products (see also AAR96021-22 and AAR96024) can be obtd.

XX from transformed host cells and used as vaccines to protect humans

XX or animals against periodontal disease. Expression in Salmonella

XX cells allows prodn. of live vaccine. Harep1-4 can also be used

XX to detect the presence of anti-P. gingivalis antibodies and to

XX raise monoclonal antibodies for diagnostic appln.

XX Sequence 456 AA:

Query Match 100.0%; Score 108; DB 17; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNLTISKDVGTATKVKY 21
 DB 251 ALNPDNLTISKDVGTATKVKY 271

RESULT 9
 AAR96990 standard; Protein: 456 AA.

XX AAR96990:

XX 22-DEC-1998 (first entry)

XX Haemagglutinin protein hga, Harep2.

XX Haemagglutinin protein: periodontal disease; vaccine; hga.

XX Porphyromonas gingivalis.

XX Sequence 456 AA:

Query Match 100.0%; Score 108; DB 17; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNLTISKDVGTATKVKY 21
 DB 251 ALNPDNLTISKDVGTATKVKY 271

XX US5824791-A.
 PN 20-OCT-1998.
 XX
 XX 11-DEC-1995; 95US-0570311.
 PF
 XX 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 XX
 DR WPI: 1998-582627/49.
 DR N-PSDB: AAV58877.
 XX
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))
 XX
 PS Claim 1: Column 127-132; 101pp; English.
 XX
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hagg haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 CC
 SO Sequence 456 AA:
 OY 1 ALNPDNYLISKDVTGATKVKY 21
 DB 251 ALNPDNYLISKDVTGATKVKY 271
 RESULT 10
 ID AAW69491 standard; Protein; 456 AA.
 XX
 AC AAW69491:
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Haemagglutinin protein hagg, HAREP3.
 XX
 KW Haemagglutinin protein; periodontal disease; vaccine; hagg.
 XX
 OS Porphyromonas gingivalis.
 XX
 PN US5824791-A.
 XX
 PD 20-OCT-1998.
 XX
 PF 11-DEC-1995; 95US-0570311.
 XX
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;

PI Tumwasorn S;
 XX
 DR WPI: 1998-582627/49.
 DR N-PSDB: AAV58878.
 XX
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))
 XX
 PS Claim 1: Column 133-138; 101pp; English.
 XX
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hagg haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 CC
 SO Sequence 456 AA:
 OY 1 ALNPDNYLISKDVTGATKVKY 21
 DB 251 ALNPDNYLISKDVTGATKVKY 271
 RESULT 11
 ID AAR96028 standard; Protein; 1087 AA.
 XX
 AC AAR96028:
 XX
 DT 04-SEP-1996 (first entry)
 XX
 DE P. gingivalis haemagglutinin haggD.
 XX
 KW Haemagglutinin; haggD; periodontal disease; vaccine; antibody.
 XX
 OS Porphyromonas gingivalis strain FDC381.
 XX
 PN W09617936-A2.
 XX
 PD 13-JUN-1996.
 XX
 PF 11-DEC-1995; 95WO-US16108.
 XX
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 XX
 DR WPI: 1996-287181/29.
 DR N-PSDB: AAT30652.
 XX
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 XX
 PS Claim 5: Page 65-68; 153pp; English.
 XX
 CC P. gingivalis 381 haemagglutinin haggD (AAR96028) was identified as
 CC the product of a gene (AAT30652) isolated from a P. gingivalis 318
 CC genomic library. The haemagglutinin (see also AAR96032) can be obt.
 CC from transformed host cells and used as a vaccine to protect humans
 CC or animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of a live vaccine. The haemagglutinin can also
 CC be used to detect the presence of anti-P. gingivalis antibodies and
 CC to raise monoclonal antibodies for diagnostic apin.

SO Sequence 1087 AA:

Query Match 100.0%; Score 108; DB 17; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 6.4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYTGATKVKY 21
 ||||||||||||||||||
 DB 579 ALNPDNYLISKDYTGATKVKY 599

RESULT 12
 AAW69486 ID AAW69486 standard; Protein: 1087 AA.

AC AAW69486;

DT 22-DEC-1998 (first entry)

DE Haemagglutinin protein hagd.

KM Haemagglutinin protein: periodontal disease; vaccine; hagd.

OS Porphyromonas gingivalis.

XX US5824791-A.

XX 20-OCT-1998.

PF 11-DEC-1995; 95US-0570311.

PR 11-DEC-1995; 95US-0570311.

PR 08-SEP-1988; 88US-0241640.

PR 25-JAN-1991; 91US-0647119.

PR 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

PI Han N, Lantz M, Lepine G, Patti JM, Progulsk-Fox A;

PI Tummasorn S;

XX WPI: 1998-582627/49.

DR N-PSDB; AAV58873.

XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin

PT and/or protease poly(peptide(s))

PS Claim 1: Column 57-64; 101pp; English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the

CC invention. This sequence represents the hagd haemagglutinin protein. The

CC polypeptides are used to produce antibodies to organisms associated with

CC periodontal disease. The antibodies are also used in purification and

CC identification procedures. The genes and polypeptides are used as

CC vaccines against periodontal disease.

XX Sequence 1087 AA:

Query Match 100.0%; Score 108; DB 19; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 6.4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYTGATKVKY 21
 ||||||||||||||||||
 DB 579 ALNPDNYLISKDYTGATKVKY 599

RESULT 13
 AAR96032 ID AAR96032 standard; Protein: 1358 AA.

AC AAR96032;

XX 04-SEP-1996 (first entry)

DE P. gingivalis hagd haemagglutinin.

KM Haemagglutinin: hagd; periodontal disease; vaccine; antibody.

OS Porphyromonas gingivalis strain FDC381.

XX WO9617936-A2.

XX 13-JUN-1996.

PF 11-DEC-1995; 95WO-US16108.

PR 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

PI Han N, Lantz M, Lepine G, Patti JM, Progulsk-Fox A;

PI Tummasorn S;

DR WPI: 1996-287181/29.

DR N-PSDB; AAT30655.

XX Porphyromonas gingivalis genes and proteins - used in the detection

PT and vaccination against periodontal disease

PS Claim 5: Page 125-129; 153pp; English.

XX P. gingivalis 381 haemagglutinin hagd (AAR96032) was identified as

CC the product of the second open reading frame of the hagd gene

CC (AAT30655) derived from P. gingivalis 318 genomic DNA. A first

CC open reading frame coded for hagd protease (see also AAR96031).

CC The protease and haemagglutinin can be obtd. from transformed host

CC cells and used in vaccines to protect humans or animals against

CC periodontal disease. Expression in Salmonella cells allows prodn.

CC of live vaccines. The haemagglutinin and protease can also be used

CC to detect the presence of anti-P. gingivalis antibodies and to raise

CC monoclonal antibodies for diagnostic appln.

XX Sequence 1358 AA:

Query Match 100.0%; Score 108; DB 17; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYTGATKVKY 21
 ||||||||||||||||||
 DB 850 ALNPDNYLISKDYTGATKVKY 870

RESULT 14
 AAW69494 ID AAW69494 standard; Protein: 1358 AA.

AC AAW69494;

DT 22-DEC-1998 (first entry)

DE Haemagglutinin protein hagd.

XX Haemagglutinin protein: periodontal disease; vaccine; hagd.

OS Porphyromonas gingivalis.

XX US5824791-A.

XX 20-OCT-1998.

PF 11-DEC-1995; 95US-0570311.

PR 11-DEC-1995: 95US-0570311.
 PR 08-SEP-1988: 88US-0241640.
 PR 25-JAN-1991: 91US-0647119.
 PR 09-DEC-1994: 94US-0353485.

XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Frogulske-Fox A;
 PI Tumwasorn S;

XX WPI: 1998-582627/49.
 DR N-PSDB: AAV58880.

XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))

XX Claim 1: Column 145-158; 101pp; English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hagd haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

XX SQ Sequence 1358 AA;

Query Match 100.0%; Score 108; DB 19; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTGATKVKY 21
 |||
 DB 850 ALNDPNYLISKDVTGATKVKY 870

RESULT 15

AA96033
 ID AA96033 standard; Protein: 1687 AA.

AC AA96033;

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin hagd.

KW Haemagglutinin; hagd; periodontal disease; vaccine; antibody.

OS Porphyromonas gingivalis strain FDC381.

XX WO9617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995: 95WO-US16108.

XX 09-DEC-1994: 94US-0353485.

PA (UABR-) UAB RES FOUND.

PA (UYFL) UNIV FLORIDA.

PI Han N, Lantz M, Lepline G, Patti JM, Frogulske-Fox A;

DR WPI: 1996-287181/29.

DR N-PSDB: AAT30656.

PT Porphyromonas gingivalis genes and proteins - used in the detection

PS Claim 5; Page 138-143; 153pp; English.

CC P. gingivalis 381 haemagglutinin hagd (AA96033) was identified as
 CC the product of a gene (AA930656) identified in P. gingivalis 318
 CC genomic DNA. The haemagglutinin can be obtd. from transformed
 CC host cells and used as a vaccine to protect humans or animals against
 CC periodontal disease. Expression in Salmonella cells allows prodn. of
 CC a live vaccine. The haemagglutinin can also be used to detect the
 CC presence of anti-P. gingivalis antibodies and to raise monoclonal
 CC antibodies for diagnostic appln.

XX SQ Sequence 1687 AA;

Query Match 100.0%; Score 108; DB 17; Length 1687;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTGATKVKY 21
 |||
 DB 1179 ALNDPNYLISKDVTGATKVKY 1199

Search completed: June 26, 2003, 00:51:42
 Job time : 7.41177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:09 ; Search time 2.20588 Seconds
(without alignments)
200.076 Million cell updates/sec

Title: US-09-980-370-1
Perfect score: 78
Sequence: 1 ALNPDNLYISKDVTG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents-AA:
1: /cgn2_6/pdata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/pdata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/pdata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/pdata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/pdata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/pdata/1/1aa/Backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	439	2	US-08-570-311-22
2	78	100.0	450	2	US-08-570-311-16
3	78	100.0	456	2	US-08-570-311-18
4	78	100.0	456	2	US-08-570-311-20
5	78	100.0	1087	2	US-08-570-311-8
6	78	100.0	1087	2	US-08-570-311-8
7	78	100.0	1358	2	US-08-570-311-27
8	78	100.0	1687	2	US-08-570-311-29
9	78	100.0	1704	3	US-08-336-308A-10
10	78	100.0	1704	3	US-08-822-324-6
11	78	100.0	1732	4	US-09-490-931-10
12	78	100.0	1732	2	US-08-570-311-10
13	78	100.0	1732	2	US-08-353-485-10
14	78	100.0	2628	2	US-08-570-311-14
15	46	59.0	497	2	US-08-570-311-2
16	46	59.0	497	2	US-08-353-485-2
17	42	53.8	2548	4	US-09-172-422-1
18	40	51.3	248	1	US-08-152-019A-14
19	40	51.3	249	1	US-08-152-019A-35
20	40	51.3	249	1	US-08-144-121-6
21	40	51.3	249	2	US-08-146-309-6
22	40	51.3	249	2	US-08-125-077-6
23	40	51.3	249	2	US-08-735-893-6
24	40	51.3	1196	1	US-08-144-121-4
25	40	51.3	1196	1	US-08-735-893-4
26	39	50.0	331	2	US-08-997-080-182
27	39	50.0	331	2	US-08-997-362-182

28	39	50.0	331	4	US-09-095-855-182	Sequence 182, App
29	39	50.0	331	4	US-09-324-542-182	Sequence 182, App
30	39	50.0	331	4	US-09-205-426-182	Sequence 182, App
31	38	48.7	638	2	US-08-846-762-95	Sequence 95, App
32	37	47.4	275	1	US-07-779-890-8	Sequence 8, App
33	37	47.4	275	1	US-07-779-890-8	Sequence 8, App
34	37	47.4	275	5	PCT-US93-05640-8	Sequence 8, App
35	37	47.4	416	4	US-09-228-246-4	Sequence 4, App
36	37	47.4	509	1	US-07-779-890-6	Sequence 6, App
37	37	47.4	509	1	US-07-779-890-6	Sequence 6, App
38	37	47.4	509	2	US-09-008-962-3	Sequence 3, App
39	37	47.4	509	2	US-08-673-307-3	Sequence 3, App
40	37	47.4	509	3	US-09-213-205-3	Sequence 3, App
41	37	47.4	509	3	US-08-733-360A-10	Sequence 10, App
42	37	47.4	509	4	US-08-916-935-11	Sequence 11, App
43	37	47.4	509	5	PCT-US93-05640-6	Sequence 6, App
44	37	47.4	524	1	US-08-447-500-24	Sequence 24, App
45	37	47.4	524	1	US-08-454-097-24	Sequence 24, App

ALIGNMENTS

RESULT 1
US-08-570-311-22
Sequence 22, Application US/08570311
Patent No. 5624791
GENERAL INFORMATION:
APPLICANT: Prologis-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patil, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: 0F15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:

LENGTH: 439 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-570-311-22

Query Match 100.0%; Score 78; DB 2; Length 439;
 Best Local Similarity 100.0%; Pred. No. 6.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDYTG 15
 DB 251 ALNPDNYLSKDYTG 265

RESULT 2
 US-08-570-311-16

Sequence 16, Application US/08570311
 Patent No. 5824791
 GENERAL INFORMATION:
 APPLICANT: Proguiske-Fox, Ann
 APPLICANT: Tumwasorn, Somying
 APPLICANT: Lepine, Guylaine
 APPLICANT: Han, Naiming
 APPLICANT: Lantz, Joseph
 APPLICANT: Patti, Joseph
 TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
 TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ted W. Whitlock
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/570,311
 FILING DATE: 25-JAN-1991
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/353,485
 FILING DATE: 09-DEC-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/647,119
 FILING DATE: 25-JAN-1991
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/241,640
 FILING DATE: 08-SEP-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Whitlock, Ted W.
 REGISTRATION NUMBER: 36,965
 REFERENCE/DOCKET NUMBER: UF15.C3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904) 375-8100
 TELEFAX: (904) 372-5800
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 450 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-570-311-16

Query Match 100.0%; Score 78; DB 2; Length 450;
 Best Local Similarity 100.0%; Pred. No. 6.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ALNPDNYLSKDYTG 15
 DB 245 ALNPDNYLSKDYTG 259

RESULT 3
 US-08-570-311-18

Sequence 18, Application US/08570311
 Patent No. 5824791
 GENERAL INFORMATION:
 APPLICANT: Proguiske-Fox, Ann
 APPLICANT: Tumwasorn, Somying
 APPLICANT: Lepine, Guylaine
 APPLICANT: Han, Naiming
 APPLICANT: Lantz, Joseph
 APPLICANT: Patti, Joseph
 TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
 TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ted W. Whitlock
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/570,311
 FILING DATE: 25-JAN-1991
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/353,485
 FILING DATE: 09-DEC-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/647,119
 FILING DATE: 25-JAN-1991
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/241,640
 FILING DATE: 08-SEP-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Whitlock, Ted W.
 REGISTRATION NUMBER: 36,965
 REFERENCE/DOCKET NUMBER: UF15.C3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904) 375-8100
 TELEFAX: (904) 372-5800
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 456 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-570-311-18

Query Match 100.0%; Score 78; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 6.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ALNPDNYLSKDYTG 15
 DB 251 ALNPDNYLSKDYTG 265
 RESULT 4

US-08-570-311-20

Sequence 20, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naïming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-20

US-08-570-311-20

Query Match 100.0%; Score 78; DB 2; Length 456;

Best Local Similarity 100.0%; Pred. No. 6.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLISKDVTG 15

DB 251 ALNPDNYLISKDVTG 265

RESULT 5

US-08-570-311-8

Sequence 8, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naïming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1087 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-8

US-08-570-311-8

Query Match 100.0%; Score 78; DB 2; Length 1087;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLISKDVTG 15

DB 579 ALNPDNYLISKDVTG 593

RESULT 6

US-08-353-485-8

Sequence 8, Application US/08353485

Patent No. 5830710

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naïming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIORITY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Query Match 100.0%; Score 78; DB 2; Length 1087;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNTLISKDVVG 15
DB 579 ALNDPNTLISKDVVG 593

RESULT 7
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Gylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

Query Match 100.0%; Score 78; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNTLISKDVVG 15
DB 850 ALNDPNTLISKDVVG 864

RESULT 8
US-08-570-311-29
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Gylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: DF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Query Match 100.0%; Score 78; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPNYLISKDVYG 15
|||||

Db 1179 ALPNPNYLISKDVYG 1193

RESULT 9
US-08-336-308A-10
Sequence 10, Application US/08336308A
Patent No. 6017532
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 100.0%; Score 78; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPNYLISKDVYG 15
|||||

Db 1196 ALPNPNYLISKDVYG 1210

RESULT 10
US-08-822-324-6
Sequence 6, Application US/08822324
Patent No. 6129917
GENERAL INFORMATION:
APPLICANT: Potempa, Jan S.
APPLICANT: Travis, James
APPLICANT: Genco, Caroline A.
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,945
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 103-95 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 488-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-324-6

Query Match 100.0%; Score 78; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPNYLISKDVYG 15
|||||

Db 1196 ALPNPNYLISKDVYG 1210

RESULT 11
US-09-490-931-10
Sequence 10, Application US/09490931
Patent No. 6274718
GENERAL INFORMATION:

APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,931
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/336,308
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-490-931-10
Query Match 100.0%; Score 78; DB 4; Length 1704;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ALNDNYLSKDVGTG 15
DB 1196 ALNDNYLSKDVGTG 1210
RESULT 12
US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiping
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10
Query Match 100.0%; Score 78; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ALNDNYLSKDVGTG 15
DB 1216 ALNDNYLSKDVGTG 1230
RESULT 13
US-08-353-485-10
Sequence 10, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiping
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485

FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-10

Query Match 100.0%; Score 78; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALNPDNYLISKDVTG 15
|||||
DB 1216 ALNPDNYLISKDVTG 1230

RESULT 14
US-08-570-311-14
Sequence 14, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patli, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Query Match 100.0%; Score 78; DB 2; Length 2628;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALNPDNYLISKDVTG 15
|||||
DB 744 ALNPDNYLISKDVTG 758

RESULT 15
US-08-570-311-2
Sequence 2, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patli, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-2

Query Match 59.0%; Score 46; DB 2; Length 497;
Best Local Similarity 61.5%; Pred. No. 3.5;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 LNPDNVLSKDYR 14
I:|||||::||
Db 291 LSPDNVLTTPKVT 303

Search completed: June 26, 2003, 00:53:58
Job time: 2.20568 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 25, 2003, 23:05:39 ; Search time 181 Seconds

(without alignments)
591.898 Million cell updates/sec

Title: US-09-980-370-5

Perfect score: 731
Sequence: 1 gcagactccacggaacglt.....ccacggaacgaatgcgcgc 402

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-O/cgn2.1/USPTO.spool/US09980370/runat.19062003.174442.8837/app_query.fasta.1.583
-DB-A.Geneseq.101002 -OPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=Diosum62 -TRA=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980370_ECGN.1.1.53_@runat.19062003.174442.8837 -NCP=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	726	99.3	134	AA849217	peptide used in th
2	716	97.9	1687	AA866033	P. gingivalis haem
3	716	97.9	1687	AA869495	Haemagglutinin pro
4	716	97.9	1704	AA870188	Arg-gingipain-2 pr
5	716	97.9	1704	AA874843	Arg-gingipain high
6	716	97.9	1704	AA874843	Arg-gingipain-2 am
7	716	97.9	1704	AA874843	P. gingivalis high
8	716	97.8	1732	AA869029	P. gingivalis high
9	715	97.8	1732	AA869029	prtk antigenic pro
10	715	97.8	1732	AA869029	Haemagglutinin pro
11	714	97.7	1706	AA874786	Haemagglutinin pro
12	711	97.3	439	AA866024	P. gingivalis haem
13	711	97.3	439	AA866024	P. gingivalis haem
14	711	97.3	450	AA866021	Haemagglutinin pro
15	711	97.3	450	AA866021	P. gingivalis haem
16	711	97.3	450	AA866021	Haemagglutinin pro
17	711	97.3	2628	AA866030	P. gingivalis haem
18	709	97.0	1087	AA866028	Haemagglutinin pro
19	709	97.0	1087	AA866028	P. gingivalis haem
20	709	97.0	1358	AA866032	Haemagglutinin pro
21	709	97.0	1358	AA866032	P. gingivalis haem
22	704	96.3	456	AA866022	Haemagglutinin pro
23	704	96.3	456	AA866022	P. gingivalis haem
24	704	96.3	456	AA866023	P. gingivalis haem
25	704	96.3	456	AA866023	Porphyromonas ging
26	205	28.0	921	AA866025	Haemagglutinin pro
27	205	28.0	922	AA866025	Porphyromonas ging
28	205	28.0	922	AA866025	Porphyromonas ging
29	205	28.0	922	AA866025	Porphyromonas ging
30	180	24.6	312	AA866025	Porphyromonas ging
31	180	24.6	312	AA866025	Porphyromonas ging
32	176.5	24.1	497	AA866025	Porphyromonas ging
33	176.5	24.1	497	AA866025	Porphyromonas ging
34	175	23.9	419	AA866025	Porphyromonas ging
35	172.5	23.6	419	AA866025	Porphyromonas ging
36	129	17.6	24	AA866025	Porphyromonas ging
37	113	15.5	20	AA866025	Porphyromonas ging
38	108	14.8	21	AA866025	Porphyromonas ging
39	95.5	13.1	196	AA866025	Porphyromonas ging
40	94	12.9	16	AA866025	Porphyromonas ging
41	90.5	12.4	285	AA866025	Porphyromonas ging
42	86.5	11.8	552	AA866025	Porphyromonas ging
43	84.5	11.6	316	AA866025	Porphyromonas ging
44	83.5	11.4	316	AA866025	Porphyromonas ging
45	83.5	11.4	316	AA866025	Porphyromonas ging

ALIGNMENTS

RESULT 1	
AA849217	
ID	AA849217 standard; protein: 134 AA.
AC	AA849217
XX	
DT	13-MAR-2001 (first entry)
XX	
DE	Peptide used in the invention.
XX	
KW	HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
XX	
OS	Infection.
XX	
PN	Unidentified.
XX	
PD	MO200072875-A1.
XX	
XX	07-DEC-2000.

Accession	Sequence	Length	Matches	Conservative	Mismatches	Indels	Gaps
PF 26-MAY-2000; 2000MO-AU00599.		7,176-79	134				
PR 28-MAY-1999; 99AU-0000652.		726-00	134				
PA (UNSV) UNIV SYDNEX.		100.00%	0				
PI Collyer CA, Hunter N, De Carlo AA;		100.00%	0				
XX WPI: 2001-080424/09.		99.32%	0				
XX		22	0				
PT Treating microbial infection in environment containing porphyrin, by							
PP administering a HA-2 antagonist							
PS Claim 9; Page 98-99; 102pp; English.							
CC The present invention relates to prophylaxis or treatment of infection							
CC by a microorganism in a biological environment comprising iron, heme							
CC or porphyrin. The treatment involves administering to the environment,							
CC an agent which antagonizes the interaction between a molecule derived							
CC from the microorganism and having an HA2 domain, and an HA2-binding							
CC motif on a porphyrin containing molecule present in the environment.							
CC useful in the manufacture of a medicament for the prophylaxis and							
CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease							
CC resulting from porphyromonas gingivalis infection or infection by a							
CC related microorganism.							
CC							
XX							
SQ Sequence 134 AA:							
Alignment Scores:							
Pred. No.:	7,176-79	Length:	134				
Score:	726.00	Matches:	134				
Percent Similarity:	100.00%	Conservative:	0				
Best Local Similarity:	100.00%	Mismatches:	0				
Query Match:	99.32%	Indels:	0				
DB:	22	Gaps:	0				
US-09-980-370-5 (1-402) x AAB49217 (1-134)							
OY 1 GCAGACTTCACGAAACGTTTCGACTCTTCTACTCATGACAGCGACCACCGAATGCACT 60							
DB 1 AIAASRPhaThcIuThrPheGluSerSerThrHisGlyIuAlaProAlaGluTrpThr 20							
OY 61 ACTATTCGATCCGCGATGCGCATGGGTGAAGGGTGGCTGCTGTCGTCCGCAATGGAC 120							
DB 21 ThrIleASpAlaAspIuAspArgIuGluGlyTrpLeuIcysLeuSerSerGlyGlnLeuAsp 40							
OY 121 TGGCTCACAGTCATGCGCGCAACCAAGTAAAGTCTTTCTCATGCAATGCAATGGCT 180							
DB 41 TrpLeuThrAlaHisGluGlyThrIleAsnValIleSerSerPheSerIlePheAsnGlyMetAla 60							
OY 181 TTGATTCCTCGATTAACATCTCTCATCTCAAAAGATGTTAAAGGATGTTAAAGTAC 240							
DB 61 LeuAsnProAspAsnIuLeuIleSerLeuAspValThrGlyAlaThrIuLysValLysTyr 80							
OY 241 TACTATTCAGTCAACAGCAGCGTTTCCCGGGGATCACTATGCGGTGATGATCTCAAGAGC 300							
DB 81 TyrTyrProValAsnAspArgIuPheProGluAspHisIuTyrluAlaValMetIleSerLysThr 100							
OY 301 GGCACGACACCGCGAGACTTTCACGGTTGTTTTCGAAGAAGCGCTTAAGCAATTAATAG 360							
DB 101 GlyThrAsnAlaGlyAspPheThrValIlePheGluGluThrProAsnGlyIleAsnLys 120							
OY 361 GCGCGAGCAAGATTCGGTCTTTCACAGGAAGCAATGGCGGC 402							
DB 121 GlyGlyAlaIa 134							
RESULT 2							
AAR96033							
XX AAR96033 standard; Protein; 1687 AA.							
XX AAR96033;							
XX							
XX 04-SEP-1996 (first entry)							

XX	P.	gingivalis	haemagglutinin	hage.
DE				
XX	Haemagglutinin;	hage;	periodontal disease;	vaccine; antibody.
KW				
XX	Porphyromonas gingivalis strain FDC381.			
OS				
XX	MO9617936-A2.			
PN				
XX	13-JUN-1996.			
PD				
XX	11-DEC-1995;	95WO-US16108.		
PF				
XX	09-DEC-1994;	94US-0353485.		
PR				
XX	(UABR-) UAB RES FOUND.			
PA	(UYFL) UNITV FLORIDA.			
FA				
XX	Han N, Lantz M, Lepine G, Patel JM, Progulske-Fox A;			
PI	Tumwasorn S;			
XX	WPI; 1996-287181/29.			
DR	N-PSDB; AAT30656.			
XX	Porphyromonas gingivalis genes and proteins - used in the detection			
PT	and vaccination against periodontal disease			
XX	Claim 5; Page 138-143; 153pp; English.			
PS				
CC	P. gingivalis 381 haemagglutinin hage (AAR96033) was identified as			
CC	the product of a gene (AAT30656) identified in P. gingivalis 318			
CC	genomic DNA. The haemagglutinin can be obdcd. from transformed			
CC	host cells and used as a vaccine to protect humans or animals against			
CC	periodontal disease. Expression in Salmonella cells allows produ. of			
CC	a live vaccine. The haemagglutinin can also be used to detect the			
CC	presence of anti-P. gingivalis antibodies and to raise monoclonal			
CC	antibodies for diagnostic appln.			
XX				
SQ	Sequence	1687 AA;		
Alignment Scores:				
	Pred. No.:	2.91e-77	Length:	1687
	Score:	716.00	Matches:	131
	Percent Similarity:	99.25%	Conservative:	1
	Best Local Similarity:	97.76%	Mismatches:	2
	Query Match:	97.95%	Indels:	0
	DB:	17	Gaps:	0
US-09-980-370-5 (1-402) x AAR96033 (1-1687)				
OY	1	GCAGACCTTCACGGAAACGTTGCAGCTTCTACTCATGAGAGGACCAACCGGAAATGACT	60	
Db	1120	AlaaSpPheHnIGuThrPhcGIUsErSerThrlHSIgYLUAAlAPRoAlaGIuTrPrNr	1139	
OY	61	ACTATGATGCCGATGCGATGCGATGAGGGTGAGGGTTCGTCTGTCTTCGGAATTGGAC	120	
Db	1140	ThrLIeaSPALlaaSPGLyaSPGLyInclYrrPlencysLeuSeSrserGLyInLenuasp	1159	
OY	121	TGGCTCACAGCTCATGGCGGACCAACAGCTAGTAAGCTTTTCTCATGGAATGGCT	180	
Db	1160	TripleuTrlraIAhISGLyGLYTThraSnaValaAlaSeSrPheSerTrIPasngILmetLIa	1179	
OY	181	TTGAATCCTGATAACTATATCTCATGTCGAAGAAGTGTTCACGGCGCAACAGGTAAAGTAC	240	
Db	1180	LeuaSnPrIoASPaSTTyrrLeuIIleSeIryLSaPYalIThrGLyAlaIThrLyValLlystYr	1199	
OY	241	TACTATCCAGTCACAGACGGTTTTTCCGGGATCACAATGCGGTGATGATTCACAAGAC	300	
Db	1200	TyrTYrrAlaValaAlaSnAsPGLYPheProGLySPHisTYrrAlaValaMetIIleSerLythr	1219	
OY	301	GGCAGCAAGCCGGAGACTTTCACGGTGTGTTTCCGAAGAAACGCCCTAACGGGAATAAATAG	360	
Db	1220	GIYTrlrAsnAlaGLyaSPpHeThrvAlaValaPhcGIUGuTrnrPrOASngLYIlEaSnLyS	1239	

CC Incorporating AG-1.

XX Sequence 1704 AA:

Alignment Scores:

Pred. No.:	2,92e-77	Length:	1704
Score:	716.00	Matches:	131
Percent Similarity:	99.25%	Conservative:	2
Best Local Similarity:	97.76%	Mismatches:	1
Query Match:	97.95%	Indels:	0
	16	Gaps:	0

US-09-980-370-5 (1-402) x AAR70188 (1-1704)

```

OY 1 GCAGACTTCACGGAAGCGTTCGATCTTCTACATGAGAGCGCACCAGCGAATGGACT 60
DB 1137 ALaAspPheThrGluThrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 1156
OY 61 ACTATCGATCCGATCGCATGCGTTCGCTGTCTGTCTTCCGACAATGGAC 120
DB 1157 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1176
OY 121 TGGCTCAGCTGATGCGCGGACCAAGCTAGTAACTTCTTCATGATGGAATGGCT 180
DB 1177 TrpLeuThrAlaHisGlyGlyThrAsnValAlaLaserPheSerTrpAsnGlyMetAla 1196
OY 181 TTGAATCCTGATTAACATCTCATCTCAAGAGATGTTACAGCGCAAGGTAAGTAC 240
DB 1197 LeuAsnProAspAsnTyrrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1216
OY 241 TACTATCCAGTCACGACGAGTTCCTCCGGGATCATTATCGGATGATCTCCAGACG 300
DB 1217 TyrTrpAlaValAlaAspGlyPheProGlyAspHisTrpAlaValMetIleSerLysThr 1236
OY 301 GGCAGCAAGCCCGAGACTTACAGCGTGTCTTTCGAAGAAGCGCTAAGCAATAATAG 360
DB 1237 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1256
OY 361 GCGGAGCAAGATTTCGCTCTTCCACGGAAGCCCAATGGCGCC 402
DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1270

```

RESULT 5

AAM34843

ID AAM34843 standard; protein: 1704 AA.

AC AAM34843:

XX 03-JUN-1998 (first entry)

XX Arg-gingipain high molecular weight prepolyprotein sequence.

XX Arg-specific gingipain protease; gingivitis; periodontal disease;

XX vaccine; infection.

XX Porphyromonas gingivalis.

XX Key Location/Qualifiers

XX Protein 1..227

XX MO9734629-A1

XX 25-SEP-1992

XX 21-MAR-1997: 97WO-US04635.

XX 22-MAR-1996: 96US-0013945.

XX (MORE-) MOREHOUSE SCHOOL MEDICINE.

XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.

XX PI Genco CA, Potempa J, Travis J, Genco C;

XX WPI: 1997-479993/44.

DR N-PSDB: AAT93872.

XX Porphyromonas gingivalis Arg-specific gingipain protease peptide(s)

PT - useful for protecting animals and humans from gingivitis and

DB periodontal diseases

XX Disclosure: Pages 68-73: 95pp: English.

The present sequence represents an arginine-specific protease of Porphyromonas gingivalis. The following peptides, derived from Arg- and Lys-specific high molecular weight proteases, offer protection against infection: YTYTVYRDG IKEGTATPE DDGATGTHH KYEKYTAG VSPKVC CC (II); YPEVEKONG RMIVYAKY (II); QLEPTDVAC YNGDFESMR CFAELKRAO CC (III); GEPNPYQPS NLATVQGR VTKMDAPSR (IV); GMEYCEVK YTAGVSPKVC CC (V); RMEFNTPEGR YTPVEKONG (VI); TFAGFEDYK RMEFNTPEGR (VII); DYTYTVYRDG TKIKELTAT TFEEDGATG NMEYCVAKY TAGVSPKVC (VIII); KIRKGLTATF EEDG (IX); RDGTIKELG TATFEEDG VTAGVSPKVC (X); KIRKGLTATF FEEDGATG HEY (XI); KMDAPNGTEN PNP NPNPN PGTTLSE (XII); and YTPVEKONG RMIVYAKY (XIII). They are used in vaccines to protect CC animals, including humans, from gingivitis and/or periodontal diseases.

XX Sequence 1704 AA:

Alignment Scores:	2,92e-77	Length:	1704
Pred. No.:	716.00	Matches:	131
Score:	99.25%	Conservative:	2
Percent Similarity:	97.76%	Mismatches:	1
Best Local Similarity:	97.95%	Indels:	0
Query Match:	97.95%	Gaps:	0

US-09-980-370-5 (1-402) x AAM34843 (1-1704)

```

OY 1 GCAGACTTCACGGAAGCGTTCGATCTTCTACATGAGAGCGCACCAGCGAATGGACT 60
DB 1137 ALaAspPheThrGluThrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 1156
OY 61 ACTATCGATCCGATCGCATGCGTTCGCTGTCTGTCTTCCGACAATGGAC 120
DB 1157 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1176
OY 121 TGGCTCAGCTGATGCGCGGACCAAGCTAGTAACTTCTTCATGATGGAATGGCT 180
DB 1177 TrpLeuThrAlaHisGlyGlyThrAsnValAlaLaserPheSerTrpAsnGlyMetAla 1196
OY 181 TTGAATCCTGATTAACATCTCATCTCAAGAGATGTTACAGCGCAAGGTAAGTAC 240
DB 1197 LeuAsnProAspAsnTyrrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1216
OY 241 TACTATCCAGTCACGACGAGTTCCTCCGGGATCATTATCGGATGATCTCCAGACG 300
DB 1217 TyrTrpAlaValAlaAspGlyPheProGlyAspHisTrpAlaValMetIleSerLysThr 1236
OY 301 GGCAGCAAGCCCGAGACTTACAGCGTGTCTTTCGAAGAAGCGCTAAGCAATAATAG 360
DB 1237 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1256
OY 361 GCGGAGCAAGATTTCGCTCTTCCACGGAAGCCCAATGGCGCC 402
DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1270

```

RESULT 6

AAT67396

ID AAT67396 standard; protein: 1704 AA.

XX AAT67396:

XX 25-APR-2000 (first entry)

XX Arg-gingipain-2 amino acid sequence.

OY		241	TACTATCATCATACAGACGGTTTCCCGGGAATCACTTCGCGTGATGTCAAAGACG	300
XX				
OY		1217	TyTtYtAlaValaAsnSpolyPheProClyAaRpnISyYalAlaValMeIleSerLyStr	1236
OY		301	GGCACGAAGCCGGAGACCTTCAGCGTTGTTTGCAGAAAGCCGCTAACGCATAAATTAG	360
Dh		1237	GlyThrsInMaIcLYaSPHeThVaIvaIPheGIuGIuThrPrOaAnGLyLeasnLys	1256
OY		361	GGCGGACCAAGATTGCGTCTTTCCACGGAAGCCATGCGCGC	402
XX				
Dh		1257	GlyGlYAlaArGPHeLIyLeuSeLTThGLuaLaSnGLYaLa	1270
RESULT 7				
AU08938	ID	AU08938 standard; protein; 1704 AA.		
AC		AAU08938;		
DT		18-DEC-2001 (first entry)		
DE		P. gingivalis hgh molecular weight Arg-gingipain-2.		
KM		Periodontitis; antiinflammatory; Arg-gingipain-2; Ag-2; immunogen.		
OS		Porphyromonas gingivalis.		
FH		Key	Location/Qualifiers	
FT		Peptide	1..227	
FT		Region	/label= Prepro-peptide	
FT		Cleavage-site	670..674	
FT		Protein	/label= Proteolytic-component	
FT		Region	719	
FT		Cleavage-site	720..1091	
FT		Protein	/label= HGP_44kDa	
FT		Region	599..619	
FT		Cleavage-site	/note= "Haemagglutinin protein component"	
FT		Protein	1091	
FT		Cleavage-site	/note= "region of homology with cysteine proteases"	
FT		Protein	1092..1429	
FT		Cleavage-site	/label= HGP_17kDa	
FT		Protein	1429	
FT		Cleavage-site	/note= "Haemagglutinin protein component"	
FT		Protein	1430..1704	
FT		Cleavage-site	/label= HGP_27kDa	
FT		Protein	/note= "Haemagglutinin protein component"	
PA		US6274718-B1.		
PD		14-AUG-2001.		
PP		25-JAN-2000; 2000US-0490931.		
PR		24-JUN-1994; 94US-0265441.		
PR		08-NOV-1994; 94US-036308.		
PR		10-SEP-1993; 93US-0119361.		
PR		09-SEP-1994; 94WO-US10283.		
PA		(UYGE-) UNIV GEORGIA RES FOUND INC.		
PI		Travis J, Potempa JS, Barr PJ, Pavloff N;		
DR		WPI: 2001-588904/66.		
DR		N-PDB: AAS15242.		
PT		New recombinant DNA molecule which encodes high molecular weight		
PT		(nature) Arg-gingipain protein; useful for immunisation against		
PT		inflammation and tissue damage; comprises enzymatically active protease		
PT		component and haemagglutinin component -		
PS		Claim 1: Column 29-41; 56pp: English.		
CC		The invention relates to a recombinant DNA molecule encoding high		

CC molecular weight (mature) Arg-gingipain (AG) protein, which has an
 CC enzymatically active protease component (AG-2) and a haemagglutinin
 CC component, from P. gingivalis. The nucleic acid is useful for producing
 CC mature Arg-gingipain protein. Immunogenic compositions comprising
 CC Arg-gingipain are useful for immunising animals including humans against
 CC inflammatory response and tissue damage caused by an archaebacterium
 CC Porphyromonas gingivalis, which causes progressive periodontitis.
 CC Arg-gingipain is also useful for identifying agents that modulate
 CC itself or preventing the interaction of the protease with the protein
 CC in the gingival area, such as complement factors C3 or C5. The
 CC present sequence is Arg-gingipain-2.

XX Sequence 1704 AA;

Alignment Scores:

Pred. No.: 2,926-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 22 Gaps: 0

US-09-980-370-5 (1-402) x AA08938 (1-1704)

OY 1 GCAGACTTCACGGAAGCTTCGAGTCTCTACATGAGAGGACGACGGAATGCACT 60
 |||||||
 DB 1137 Alaasphethrgrlurhrphgluserserthrhlsglylualaproalaglutprthr 1156
 OY 61 ACTATCGATGCGCATGGCCATGGTGGAGGCTGCTGCTGCTGCTGCTGCGACATTTGAC 120
 |||||||
 DB 1157 Thrileaspalaspolyaspolyglnyltrpleucysleuserserlglmeuasp 1176
 OY 121 TGGCTCAGACTGATGCGGCGGACCAAGCTAGTAAAGCTCTTTCATGATGAAATGGCT 180
 |||||||
 DB 1177 Trpleutnralahlsiglylthrashnvalalaserphesertrpasmolymetala 1196
 OY 181 TTGAATCCGATTAATCTCATCTCAATCAAGATGTTACAGCGGCAAGAGTAAATGAC 240
 |||||||
 DB 1197 Leuasnproaspasnryrleuileserlysaspvalthrclyalathrclyvallystyr 1216
 OY 241 TACTATCCAGTCAACGACGCTTTTCCCGGAGTACTACTGCGGTCATGATCTCAAGACG 300
 |||||||
 DB 1217 Tyrtyralavalasnsapglypserpocglyaspilistyrvalavalmecrliserlystyr 1236
 OY 301 GCGACGAACGCGGAGACTTACAGGCTGTTTTCGAAGAAGCGCTAACGAATTAATAG 360
 |||||||
 DB 1237 Glythrashnalaglyaspserthrvalvalphelglutlurhrproasnlylalelnlys 1256
 OY 361 GCGGAGGAGGATTCGCTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 402
 |||||||
 DB 1257 Glyglyalalarphelglyleuserthrqlualasnslyala 1270

RESULT 8
 AAR96029
 ID AAR96029 standard; Protein: 1732 AA.
 XX
 AC AAR96029;
 XX
 DT 04-SEP-1996 (first entry)
 XX
 DE P. gingivalis porphyrein.
 XX
 KW Porphyrein; haemagglutinin; periodontal disease; vaccine; antibody.
 XX
 OS Porphyromonas gingivalis strain W12.
 XX
 FI Key Location/Qualifiers
 FT Region 688..708
 FT /note="Pro-Asn repeat region type 1"
 FT Region 887..952
 FT /note="Pro-Asn repeat region type 2"
 FT Region 946..967

FT
 FT Region
 FT 985..1006
 FT /note="Pro-Asn repeat region type 3"
 FT Region
 FT 1041..1100
 FT /note="Pro-Asn repeat region type 4"
 FT Region
 FT 1341..1405
 FT /note="Pro-Asn repeat region type 2"
 FT Region
 FT 1430..1451
 FT /note="Pro-Asn repeat region type 3"
 FT Region
 FT 1488..1547
 FT /note="Pro-Asn repeat region type 4"
 FT Region
 FT 1607..1650
 FT /note="Pro-Asn repeat region type 2"
 XX
 EN W09617936-A2.
 XX
 PD 13-JUN-1996.
 XX
 PF 11-DEC-1995; 95MO-US16108.
 XX
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES. FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N. Lantz M, Lepine G, Patti JM, Prognuske-fox A;
 PI Tumasorn S;
 XX
 DR WPI: 1996-287181/29.
 DR N-PSDB; AAT30653.
 XX
 FT Porphyromonas gingivalis genes and proteins - used in the detection
 FT and vaccination against periodontal disease
 XX
 PS Claim 5; Page 76-81; 153pp: English.
 XX
 CC P. gingivalis W12 cysteine protease, porphyrein (AAR96029), was
 CC identified as the product of the prt gene (AAT30653) isolated from
 CC P. gingivalis W12 genomic DNA. The porphyrein shows homology to
 CC the haemagglutinins (see also AAR96026-28 and AAR96030-33) of P.
 CC gingivalis 318. It can be obtained from transformed host cells and
 CC used as a vaccine to protect humans or animals against periodontal
 CC disease. Expression in salmonella cells allows production of a live
 CC vaccine. The porphyrein and haemagglutinins can also be used to
 CC detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic application.
 XX
 SO Sequence 1732 AA;
 XX
 Alignment Scores:
 Pred. No.: 3,886-77 Length: 1732
 Score: 715.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.81% Indels: 0
 DB: 17 Gaps: 0

US-09-980-370-5 (1-402) x AAR96029 (1-1732)

OY 1 GCAGACTTCACGGAAGCTTCGAGTCTCTACATGAGAGGACGACGGAATGCACT 60
 |||||||
 DB 1157 Alaasphethrgrlurhrphgluserserthrhlsglylualaproalaglutprthr 1176
 OY 61 ACTATCGATGCGCATGGCCATGGTGGAGGCTGCTGCTGCTGCTGCTGCGACATTTGAC 120
 |||||||
 DB 1177 Thrileaspalaspolyaspolyglnyltrpleucysleuserserlglmeuasp 1196
 OY 121 TGGCTCAGACTGATGCGGCGGACCAAGCTAGTAAAGCTCTTTCATGATGAAATGGCT 180
 |||||||
 DB 1197 Trpleutnralahlsiglylthrashnvalalaserphesertrpasmolymetala 1216
 OY 181 TTGAATCCGATTAATCTCATCTCAATCAAGATGTTACAGCGGCAAGAGTAAATGAC 240
 |||||||

Db	1217	Leuansnrfroaspmsrntylleuileserlyaspvalthngilyalatrhlsvallstyr	1236
Oy	241	TACTATCACTACGACGACGGTTTCCCGGGATCACTATGGGGATGATCTCCAGACG	300
Db	1237	tyrrlravlaiaaspolyphenrogllyasphstlyralavalmelileserlysthr	1236
Oy	301	GGCAGAACCCCGGACACTTCACGGTTGTGTTTGAGAAAGCCCTAACGGATMAATAAG	360
Db	1257	Glythrsmnlaolylaspherthrvlaivalphaglucglutthrvocanngilyleasnlys	1276
Oy	361	GGCGAGACAGATTGGCTGCTTCCACGGAAGCCCATGGCCG	402
Db	1277	Glycylvalarphedglyleuserthrculalasnnglyala	1290
RESULT 9			
AAW24787	ID	AAW24787 standard; Protein; 1732 AA.	
XX	AAW24787:		
XX	25-NOV-1997	(first entry)	
XX	DE	PrTx antigenic protein complex.	
XX	XX		
XX	XX	Periodontal disease; cell surface protein; thiol protease;	
XX	XX	endopeptidase; PrTx; PrTx48; PrTx39; PrTx15; PrTx44;	
XX	XX	hemagglutinin; adhesin; therapy; diagnosis; vaccine; antigen.	
XX	OS	Porphyromonas gingivalis strain W50.	
XX	XX		
XX	XX	Key	Location/Qualifiers
XX	XX	Peptide	1..228
XX	XX		/label= Pro-pro-peptide
XX	XX	Cleavage-site	228..229
XX	XX		/note= 229..737
XX	XX	Protein	737..738
XX	XX		/label= PrTx48
XX	XX	Cleavage-site	737..738
XX	XX		/note= "48 kDa Lys-specific thiol protease"
XX	XX	Protein	738..1156
XX	XX		/label= PrTx39
XX	XX	Cleavage-site	1156..1157
XX	XX		/note= "39 kDa adhesin"
XX	XX	Protein	1157..1291
XX	XX		/label= PrTx15
XX	XX	Cleavage-site	1291..1292
XX	XX		/note= "15 kDa adhesin"
XX	XX	Protein	1292..1732
XX	XX		/label= PrTx44
XX	XX		/note= "44 kDa adhesin"
XX	XX		
XX	XX	MO9716542-A1.	
XX	XX		
XX	XX	09-MAY-1997.	
XX	XX		
XX	XX	30-OCT-1996;	96WO-AU00673.
XX	XX		
XX	XX	30-OCT-1995;	95AU-0006275.
XX	XX		
XX	XX	(UYME) UNIV MELBOURNE.	
XX	XX	(VICT-) VICTORIAN DAIRY IND AUTHORITY.	
XX	XX		
XX	XX	Bhogal PS, Reynolds EC, Slakaski N;	
XX	XX		
XX	XX	WPI; 1997-272112/24.	
XX	XX		
XX	XX	N-PSDB; AAT78851.	
XX	XX		
XX	XX	New antigenic protein complex from Porphyromonas gingivalis -	
XX	XX	comprising Arg- and Lys- specific thiol endo-peptidase(s), used in	
XX	XX	the detection, prevention and treatment of periodontal disease	
XX	XX		
XX	XX	Example 1; Fig 9b; 68pp; English.	
XX	XX		
CC	A	PrTx-PrTx cell surface protein of Porphyromonas gingivalis (PG)	

	CC	comprised a 300 kDa complex composed of a 48 kDa lysine-specific
	CC	thiol protease and 39, 15 and 44 kDa adhesins encoded by the prtK
	CC	gene (AAT7885), and a 45 kDa arginine-specific thiol protease and
	CC	44, 15, 17 and 27 kDa adhesins (see AAM24786) encoded by the prtR
	CC	gene (AAT78850). A claimed antigenic complex comprises at least one
	CC	multimeric protein complex of prtR and prtK each containing at
	CC	least one adhesion domain, the complex having a mol.wt. of over 200
	CC	kDa, and preferably comprises all 9 proteins of the PrtR-PrtK
	CC	complex (see also AAM24780-85). It can be used in a claimed
	CC	composition to elicit an immune response directed against PG, and/or
	CC	in a claimed method of reducing the prospect of PG infection and/or
	CC	severity of disease. Antibodies directed against the complex are
	CC	claimed for use in treating PG infection. Unlike whole PG cells or
	CC	other previously prepared antigens based on fibrillae or the
	CC	capsule, the PrtR-PrtK complex or component parts are safe and
	CC	effective antigens.
XX		
SQ	Sequence	1732 AA:
	Alignment Scores:	
Pred. No.:	3.88e-77	Length: 1732
Score:	715.00	Matches: 121
Percent Similarity:	99.25%	Conservative: 2
Best Local Similarity:	97.76%	Mismatches: 1
Query Match:	97.81%	Indels: 0
DB:	18	Gaps: 0
	US-09-980-370-5 (1-402) x AAM24787 (1-1732)	
OY	1 GCAGACTTCACGGAAGAAGCTTGCGATTCTTCACTCATGAGAGGCACCAGCGGAATGGACT	60
Db	1157 AIAASpPheTrHrGIuThrPhgGIuSerSerThHISGLuAlaIProAlaGIuTPPrHr	1176
OY	61 ACTATCGATWGCCGATGGCGATGGTAGGGTTGGCTGTCTCTTCGCGACAATTGGAC	120
Db	1177 ThrILespaLlaSpGIaSpGIyGIInclYlrPneucylLeuSerSercylGIInleuaSp	1196
OY	121 TGGCPCACAGCTCATGGCGCGAACAACTAAAGCTTTCTCATGGTAAGCAATGGCT	180
Db	1197 TrPbuThrAlaHISGLySerSnaValaISerSerPheSerTrIPasInclYMeLaIs	1215
OY	181 TTGAATCCTGATAACTATCTCATCTCAAAGGATGTTACAGSGCGAACGAGTTAAGTAC	240
Db	1217 LeuAsnProAbSpSnryrLeuIIeSerLySaPvalThrGlYAlaThrLySvalLyStyr	1236
OY	241 TACTCTCCAGTCAACGACGCGTTTTTCGCGGATCAGTATGCGGTGATGATCTCCAGACG	300
Db	1237 TyfTrfAlaVaLIaAsnaspGIyPheProLIySpHisTrfAlaValMeIIeSerLySrHr	1256
OY	301 GGCAGAACGCCGGAACACTTCACGGTGTTTTTCGAAGAAAGCCCTAACGAAATAATAG	360
Db	1257 GlyTHraSnAlaIGlySaPheThrValaIPheLIuIduIuThrProasnclYIleasnLyS	1276
OY	361 GGCGAGCAAGATTGCGTCTTTCACGGAAGCAAGCGCGCC	402
Db	1277 GlyGIyAlaIarGPheGIyLeuSerThrLuIluaInclYAla	1290
RESULT 10		
AAM69487	ID	AAM69487 standard: Protein; 1732 AA.
XX		
AC	AAM69487:	
XX		
DT	22-DEC-1998	(first entry)
XX		
DE	Haemagglutinin protein prtf.	
XX		
KM	Haemagglutinin protein; periodontal disease: vaccine: prtf.	
OS	Porphyromonas gingivalis.	
XX		
FN	US5824791-A.	

PD 20-OCT-1998.
 XX 11-DEC-1995: 95US-0570311.
 XX 11-DEC-1995: 95US-0570311.
 PR 08-SEP-1988: 88US-0241640.
 PR 25-JAN-1991: 91US-0547119.
 PR 09-DEC-1994: 94US-0353485.
 XX (UABR-) UAB RES FOUND.
 PA (UFL) UNIV FLORIDA.
 XX Han N, Lantz M, Lepline G, Patli JM, Proguiske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB: AAV58874.
 XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT end/or protease poly(peptide(s))
 PS Claim 1: Column 69-84: 101bp: English.
 XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the prp haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 SO Sequence 1732 AA:
 Alignment Scores:
 Pred. No.: 3,88e-77 Length: 1732
 Score: 715.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.81% Indels: 0
 DB: Gaps: 0
 US-09-980-370-5 (1-402) x AAM69487 (1-1732)
 OY 1 GCAGACTTCACGGAAGGTTGCGATCTACTCATGAGGACCGCGGATGACG 60
 DB 1157 AAlaSPRHeThrGluThrPheGluSerSerThrHisGluAlaProAlaGluTyrThr 1176
 OY 61 ACTATCGATGCCGATGGCGATGGAGGGTTGGCTGTCTGTCTCCGACAAATGGAC 120
 DB 1177 ThrIleAspAlaAspGluAspGluGlnGlyTyrPheCysLeuSerSerGlyGlnLeuAsp 1196
 OY 121 TGGCTCAGACGCTCATGGCGGCGACCAAGCTAGTAACTCTTCTCATGGAATGAGTGGCT 180
 DB 1197 TrpLeuThrAlaHisGlyGlySerAspValValSerSerPheSerTrpAsnGlyMetAla 1216
 OY 181 TGGAACTCTGATATCTATCTATCTCAAGAGTATGAGGGCGACCAAGGTAATATGC 240
 DB 1217 LeuAsnProAspAsnTyrLeuIleSerTyrAspValThrGlyAlaThrLysValLysTyr 1236
 OY 241 TACTATCCATCAACGACGCGTTTCCCGGATCAGTATGCGGATGATGATCTCCAGACG 300
 DB 1237 TyrTyrAlaValAlaAspGluPheProGlyAspHisTyrAlaValMetIleSerLysThr 1256
 OY 301 GGCAGCAACCGCGGAGACTTCACGCGTGTGTTTTCGAAGAACCGCTTAACGGAATAAATAG 360
 DB 1257 GlyThrAsnAlaGlnGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1276
 OY 361 GCGGAGCAGAGATTCGCTCTTCCACGGAACCAATGGCGCC 402
 DB 1277 GlyGlyAlaHisArgPheGlyLeuSerThrGlyAlaAsnGlyAla 1290
 RESULT 11
 AAM24786
 ID AAM24786 standard: Protein: 1706 AA.

XX AAM24786:
 AC 25-NOV-1997 (first entry)
 DT PrtR antigenic protein complex.
 XX Periodontal disease; cell surface protein; thiol protease;
 DE endopeptidase; PrtR; PrtR45; PrtR44; PrtR15; PrtR17; PrtR27;
 XX haemagglutinin; adhesion; therapy; diagnosis; vaccine; antigen.
 XX Porphyromonas gingivalis strain W50.
 OS
 XX Key
 FH Peptide
 FH Location/Qualifiers
 FT 1..227
 FT /label= Pro-pro-peptide
 FT 227..228
 FT /label= PrtR45
 FT 228..719
 FT /label= PrtR45
 FT /note= "45 kDa Arg-specific thiol protease"
 FT 719..720
 FT /label= PrtR44
 FT 720..1138
 FT /label= PrtR44
 FT /note= "44 kDa adhesin"
 FT 1138..1139
 FT /label= PrtR15
 FT 1139..1273
 FT /label= PrtR15
 FT /note= "15 kDa adhesin"
 FT 1273..1274
 FT /label= PrtR17
 FT 1274..1431
 FT /label= PrtR17
 FT /note= "17 kDa adhesin"
 FT 1431..1432
 FT /label= PrtR27
 FT 1432..1706
 FT /label= PrtR27
 FT /note= "27 kDa adhesin"
 MO9716542-AI.
 PD 09-MAY-1997.
 PD 30-OCT-1996: 96MO-AU00673.
 PR 30-OCT-1995: 95AU-0006275.
 PA (UYME) UNIV MELBOURNE.
 PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 PI Bhogal PS, Reynolds EC, Slakeski N;
 DR WPI: 1997-272112/24.
 DR N-PSDB: AAT78850.
 PT New antigenic protein complex from Porphyromonas gingivalis -
 PT comprising Arg- and Lys- specific thiol endopeptidase(s) used in
 PT the detection, prevention and treatment of periodontal disease
 XX
 PS Example 1: Fig 8b: 68pp: English.
 CC A PrtR-prtK cell surface protein of Porphyromonas gingivalis (PG)
 CC comprises a 300 kDa complex composed of a 45 kDa arginine-specific
 CC thiol protease and 44, 15, 17 and 27 kDa adhesins encoded by the
 CC prtR gene (AAT78850), and a 148 kDa lysine-specific thiol protease
 CC and 35, 15 and 44 kDa adhesins (see AAM24787) encoded by the prtK
 CC gene (AAT78851). A claimed antigenic complex comprises at least one
 CC multimeric protein complex of PrtR and PrtK each containing at
 CC least one adhesion domain, the complex having a mol.wt. of over 200
 CC kDa, and preferably comprises all 9 proteins of the PrtR-prtK
 CC complex (see also AAM24780-85). It can be used in a claimed
 CC composition to elicit an immune response directed against PG, and
 CC in a claimed method of reducing the prospect of PG infection and/or
 CC severity of disease. Antibodies directed against the complex are
 CC claimed for use in treating PG infection. Unlike whole PG cells or

CC other previously prepared antigens based on fimbriae or the
CC capsule, the PrtR-PrtK complex or component parts are safe and
CC effective antigens.

XX Sequence 1706 AA;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:
Score:	5.09e-77	1706	131	
Percent Similarity:	714.00			
Best Local Similarity:	99.25%			
Query Match:	97.76%			
DB:	18	Gaps:	0	

US-09-980-370-5 (1-402) x AAM24786 (1-1706)

OY 1 GCAGACTTCACGGAAGCTTCGAGTCTCTACATGAGAGGACCCAGCGAATGGACT 60

DB 1139 AlasprhethrlnrPheluserSerThrhlsGlyAlaProAlaGlturPrthr 1158

OY 61 ACTATCGATCCGATGCGGATGTGAGGGTTCGCTCTGTCTGTCGCGACATTTGGAC 120

DB 1159 ThrileasprAlaAspGlyAspGlyGlnGlyTrpleucysleuserSerGlyGlnleuAsp 1178

OY 121 TGCGTCACAGCTCATGGCGGCACACAGTAAAGCTTTCTCATGGAATGGAATGGCT 180

DB 1179 TrpleuthrAlaHlsGlyGlyThrAsnValAlaSerSerPheSerTrpAsnGlyMetAla 1198

OY 181 TTGAATCCTGATTAATCTATCTCAAGAGATGTACAGCGGACGACGAAAGTAAAGTAC 240

DB 1199 LeuAsnProAspAsnTyrlleuIleSerLysAspAlaThrhGlyAlaThrLysValLysTyr 1218

OY 241 TACTATCCAGTCAACGAGGTTTCCGCGGATCACTATGCGGATGATCTCCAGACG 300

DB 1219 TyrlGlyAlaValAlaAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1238

OY 301 GGCAGCAAGCGCGGACACTTACGCGTGTTCGAGAAGACGCTTAACGGAATTAATAG 360

DB 1239 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 1258

OY 361 GCGGAGCAAGATTCGCTCTTCACGAGCAAGCCCAATGGCGGC 402

DB 1259 GlyGlyAlaAspPheGlyLeuSerThrhGluAlaAspGlyAla 1272

RESULT 12
AAR96024 standard; Protein: 439 AA.

XX AAR96024:

DT 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin *haga* Harep4 product.

DE Haemagglutinin; *haga*; periodontal disease; vaccine; antibody;

KW Harep4.

OS Porphyromonas gingivalis strain 381.

XX MO9617936-A2.

PD 13-JUN-1996.

XX 11-DEC-1995; 95WO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patli JM, Progulsk-Fox A;
PI Tumworn S;
XX

DR WPI: 1996-287181/29;
DR N-PSDB: AAT30648.

PT Porphyromonas gingivalis genes and proteins - used in the detection
PT and vaccination against periodontal disease

PS Claim 4; Page 114-115; 153pp; English.

XX Harep4 (AAR96024) is the product of the Harep4 repeat unit (AAT30648)

CC of the *haga* gene (AAT30648) of *P. gingivalis* 318. It forms part

CC of haemagglutinin *haga* (see also AAR96030). Harep4 and other

CC *haga* repeat unit products (see also AAR96021-23) can be obtained from

CC transformed host cells and used as vaccines to protect humans or

CC animals against periodontal disease. Expression in *Salmonella*

CC cells allows production of live vaccine. Harep4-4 can also be used

CC to detect the presence of anti-*P. gingivalis* antibodies and to

CC raise monoclonal antibodies for diagnostic application.

XX Sequence 439 AA;

US-09-980-370-5 (1-402) x AAR96024 (1-439)

OY 1 GCAGACTTCACGGAAGCTTCGAGTCTCTACATGAGAGGACCCAGCGAATGGACT 60

DB 192 AlasprhethrlnrPheluserSerThrhlsGlyAlaProAlaGlturPrthr 211

OY 61 ACTATCGATCCGATGCGGATGTGAGGGTTCGCTCTGTCTTCGCGACATTTGGAC 120

DB 212 ThrileasprAlaAspGlyAspGlyGlnGlyTrpleucysleuserSerGlyGlnleuAsp 231

OY 121 TGCGTCACAGCTCATGGCGGCACACAGTAAAGCTTTCTCATGGAATGGAATGGCT 180

DB 232 TrpleuthrAlaHlsGlyGlyThrAsnValAlaSerSerPheSerTrpAsnGlyMetAla 251

OY 181 TTGAATCCTGATTAATCTATCTCAAGAGATGTACAGCGGACGACGAAAGTAAAGTAC 240

DB 252 LeuAsnProAspAsnTyrlleuIleSerLysAspAlaThrhGlyAlaThrLysValLysTyr 271

OY 241 TACTATCCAGTCAACGAGGTTTCCGCGGATCACTATGCGGATGATCTCCAGACG 300

DB 272 TyrlGlyAlaValAlaAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 291

OY 301 GGCAGCAAGCGCGGACACTTACGCGTGTTCGAGAAGACGCTTAACGGAATTAATAG 360

DB 292 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 311

OY 361 GCGGAGCAAGATTCGCTCTTCACGAGCAAGCCCAATGGCGGC 402

DB 312 GlyGlyAlaAspPheGlyLeuSerThrhGluAlaAspGlyAla 325

RESULT 13

XX AAM69492 standard; Protein: 439 AA.

XX AAM69492:

DT 22-DEC-1998 (first entry)

XX Haemagglutinin protein *haga*, Harep4.

XX Haemagglutinin protein; periodontal disease; vaccine; *haga*.

XX Porphyromonas gingivalis.
OS
XX US5824791-A.

XX 20-OCT-1998.
PD
XX
XX 11-DEC-1995; 95US-0570311.
PF
XX 11-DEC-1995; 95US-0570311.
PR
PR 08-SEP-1988; 88US-0241640.
PR 25-JAN-1991; 91US-0647119.
PR 09-DEC-1994; 94US-0353485.
XX
XX (UABR-) UAB RES FOUND.
PA (UYFL) UNIV FLORIDA.
XX
XX Han N, Lantz M, Lepine G, Patti JM, Progulske-Fox A;
PI Tumasorn S;
XX
XX WPI; 1998-582627/49.
DR N-PSDB; AAV58879.
XX
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
PT and/or protease poly:peptide(s))
XX
XX Claim 1; Column 139-144; 101pp; English.
XX
XX This sequence is encoded by a Porphyromonas gingivalis gene of the
CC invention. This sequence represents the hga haemagglutinin protein. The
CC polypeptides are used to produce antibodies to organisms associated with
CC periodontal disease. The antibodies are also used in purification and
CC identification procedures. The genes and polypeptides are used as
CC vaccines against periodontal disease.
XX
SQ Sequence 439 AA;

Alignment Scores:
Pred. No.: 7,15e-77 Length: 439
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 19 Gaps: 0

US-09-980-370-5 (1-402) x AAV69492 (1-439)

OY 1 GCAGACTTCACGGAACGTTGCGAGTCTTCTACTCATGAGAGGACACCGGATGACT 60
DB 192 AlaAspRheThcIuThrPheGluSerSerThrHisGlyAlaProIaGluTrpThr 211
OY 61 ACTATCGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 120
DB 212 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 231
OY 121 TGGGTCACAGCTCATGCGGCGGACCAAGTAGTAAAGCTTTTCTCATGGAATGAGTGC 180
DB 222 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 251
OY 181 TTGAATCCGTAACTATCTCATCTCAAGAGATGTTACAGGCGCAACGAAAGTAAAGTAC 240
DB 252 LeuAsnProAspAsnTyIleuIleSerGlyAspValThrGlyAlaThrIlyValIlystyr 271
OY 241 TACTATCCAGTCAACGAGCGGTTTTCCGGGATCACTATGCGGATGATGATGATGATG 300
DB 272 TyIleAlaValAlaAspGlyAspGlyAspGlyAspGlyAspGlyAspGlyAspGly 291
OY 301 GGCAGAGCGCGGAGACTTCAGGTTCTTTCGAAGAAGCCGTAAGCGATTAATAG 360
DB 292 GlyThrAsnAlaGlyAspRheThrValAlaIleGlnGlyIleThrProAsnGlyIleAsnLys 311
OY 361 GGCAGAGCAAGATGCGTCTTTCACGAGCAAGCAATGCGGC 402
DB 312 GlyGlyAlaIleArgPheGlyLeuSerThrGlyAlaAspGlyAla 325
RESULT 14
AA96021

ID AAR96021 standard; Protein; 450 AA.
XX
XX AAR96021:
XX
XX 04-SEP-1996 (first entry)
XX
XX P. gingivalis haemagglutinin hga Harepl product.
DE
XX
XX Haemagglutinin; hga; periodontal disease; vaccine; antibody;
KW Harepl.
XX
XX OS Porphyromonas gingivalis strain 381.
XX
XX WO9617936-A2.
XX
XX 13-JUN-1996.
XX
XX 11-DEC-1995; 95WO-US16108.
XX
XX 09-DEC-1994; 94US-0353485.
XX
XX (UABR-) UAB RES FOUND.
PA (UYFL) UNIV FLORIDA.
XX
XX Han N, Lantz M, Lepine G, Patti JM, Progulske-Fox A;
PI Tumasorn S;
XX
XX WPI; 1996-287181/29.
DR N-PSDB; AAT30645.
XX
XX Porphyromonas gingivalis genes and proteins - used in the detection
PT and vaccination against periodontal disease
XX
XX Claim 4; Page 103-104; 153pp; English.
XX
XX Harepl (AAR96021) is the product of the Harepl repeat unit (AAT30645)
CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
CC of haemagglutinin hga (see also AAR96030). Harepl and other
CC hga repeat unit products (see also AAR96022-24) can be obtd. from
CC transformed host cells and used as vaccines to protect humans or
CC animals against periodontal disease. Expression in Salmonella
CC cells allows prodn. of live vaccine. Harepl-4 can also be used
CC to detect the presence of anti-P. gingivalis antibodies and to
CC raise monoclonal antibodies for diagnostic appln.
XX
SQ Sequence 450 AA;

Alignment Scores:

Pred. No.: 7,22e-77 Length: 450
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 17 Gaps: 0

US-09-980-370-5 (1-402) x AAR96021 (1-450)

OY 1 GCAGACTTCACGGAACGTTGCGAGTCTTCTACTCATGAGAGGACACCGGATGACT 60
DB 186 AlaAspRheThcIuThrPheGluSerSerThrHisGlyAlaProIaGluTrpThr 205
OY 61 ACTATCGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 120
DB 206 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 225
OY 121 TGGGTCACAGCTCATGCGGCGGACCAAGTAGTAAAGCTTTTCTCATGGAATGAGTGC 180
DB 226 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 245
OY 181 TTGAATCCGTAACTATCTCATCTCAAGAGATGTTACAGGCGCAACGAAAGTAAAGTAC 240
DB 246 LeuAsnProAspAsnTyIleuIleSerGlyAspValThrGlyAlaThrIlyValIlystyr 265

•
1
-
•

Thu Jun 26 11:58:01 2003

us-09-980-370-1.rag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:59 ; Search time 5.29412 Seconds
(without alignments)
377.543 Million cell updates/sec

Title: US-09-980-370-1

Sequence: 1 ALNPDNLTISKDVTS 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.101002.*
1: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgcdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	78	100.0	15	AA849215	Residues 1215-1229
2	78	100.0	21	AA849218	Peptide #3 unde
3	78	100.0	134	AA849217	Peptide used in th
4	78	100.0	439	AA849214	P. gingivalis haem
5	78	100.0	439	AA849219	Haemagglutinin pro
6	78	100.0	450	AA849221	P. gingivalis haem
7	78	100.0	450	AA849222	Haemagglutinin pro
8	78	100.0	456	AA849223	P. gingivalis haem
9	78	100.0	456	AA849224	P. gingivalis haem
10	78	100.0	456	AA849225	Haemagglutinin pro

11	78	100.0	456	19	AA849215	Haemagglutinin pro
12	78	100.0	1087	17	AA849218	P. gingivalis haem
13	78	100.0	1087	19	AA849217	Haemagglutinin pro
14	78	100.0	1358	17	AA849214	P. gingivalis haem
15	78	100.0	1358	19	AA849219	Haemagglutinin pro
16	78	100.0	1687	17	AA849221	P. gingivalis haem
17	78	100.0	1687	19	AA849222	Haemagglutinin pro
18	78	100.0	1704	16	AA849223	P. gingivalis haem
19	78	100.0	1704	18	AA849224	P. gingivalis haem
20	78	100.0	1704	21	AA849225	Haemagglutinin pro
21	78	100.0	1704	22	AA849226	Arg-gingipain-2 pr
22	78	100.0	1704	22	AA849227	Arg-gingipain-2 am
23	78	100.0	1732	17	AA849228	P. gingivalis haem
24	78	100.0	1732	18	AA849229	P. gingivalis haem
25	78	100.0	1732	18	AA849230	P. gingivalis haem
26	78	100.0	1732	18	AA849231	P. gingivalis haem
27	78	100.0	1732	18	AA849232	P. gingivalis haem
28	78	100.0	1732	18	AA849233	P. gingivalis haem
29	78	100.0	1732	18	AA849234	P. gingivalis haem
30	78	100.0	1732	18	AA849235	P. gingivalis haem
31	78	100.0	1732	18	AA849236	P. gingivalis haem
32	78	100.0	1732	18	AA849237	P. gingivalis haem
33	78	100.0	1732	18	AA849238	P. gingivalis haem
34	78	100.0	1732	18	AA849239	P. gingivalis haem
35	78	100.0	1732	18	AA849240	P. gingivalis haem
36	78	100.0	1732	18	AA849241	P. gingivalis haem
37	78	100.0	1732	18	AA849242	P. gingivalis haem
38	78	100.0	1732	18	AA849243	P. gingivalis haem
39	78	100.0	1732	18	AA849244	P. gingivalis haem
40	78	100.0	1732	18	AA849245	P. gingivalis haem
41	78	100.0	1732	18	AA849246	P. gingivalis haem
42	78	100.0	1732	18	AA849247	P. gingivalis haem
43	78	100.0	1732	18	AA849248	P. gingivalis haem
44	78	100.0	1732	18	AA849249	P. gingivalis haem
45	78	100.0	1732	18	AA849250	P. gingivalis haem

ALIGNMENTS

RESULT 1
ID AA849215 standard; peptide: 15 AA.
AC AA849215;
DT 13-MAR-2001 (first entry)
DE Residues 1215-1229 of HA2 gingipain domain.
KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
infection.
OS Porphyromonas gingivalis.
XX WO200072875-A1.
XX 07-DEC-2009.
XX 26-MAY-2000; 2000MO-AU00599.
XX 28-MAY-1999; 99AU-0000652.
XX (UNSY) UNTV SYDNEY.
XX COLLYER CA, Hunter N, De Carlo AA.
XX WPI, 2001-080424/09.
XX Treating microbial infection in environment containing porphyrin, by
XX administering a HA-2 antagonist
XX Claim 11: Page 44; 102pp; English.
PS

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 15 AA:

Query Match 100.0%; Score 78; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVTG 15
 DB 1 ALNDPNYLISKDVTG 15

RESULT 2

AA049218 ID A049218 standard; peptide: 21 AA.

XX A049218:

DT 13-MAR-2001 (first entry)

DE Peptide #3.

XX HA2 domain: porphyrin: periodontal; pulmonary; vaginal; urethral;
 KM infection.

OS Unidentified.

PN W0200072875-A1.

PD 07-DEC-2000.

FE 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNIV SYDNEY.

XX Collyer CA, Hunter N, De Carlo AA;

WI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist
 XX Claim 11: Page 40; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 21 AA:

Query Match 100.0%; Score 78; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVTG 15
 DB 1 ALNDPNYLISKDVTG 15

RESULT 3

AA049217 ID A049217 standard; protein: 134 AA.

XX A049217:

DT 13-MAR-2001 (first entry)

DE Peptide used in the invention.

XX HA2 domain: porphyrin: periodontal; pulmonary; vaginal; urethral;
 KM infection.

OS Unidentified.

PN W0200072875-A1.

PD 07-DEC-2000.

FE 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNIV SYDNEY.

XX Collyer CA, Hunter N, De Carlo AA;

WI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist
 XX Claim 9; Page 98-99; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 134 AA:

Query Match 100.0%; Score 78; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVTG 15
 DB 60 ALNDPNYLISKDVTG 74

RESULT 4

AA06024 ID A06024 standard; protein: 439 AA.

XX A06024:

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin haga Harep4 product.

KM Haemagglutinin: haga: periodontal disease; vaccine; antibody;

KM HAREP4.
 XX Porphyromonas gingivalis strain 381.
 OS
 XX MO9617936-A2.
 PM
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995: 95WO-US16108.
 PF
 XX 09-DEC-1994: 94US-0353485.
 PR
 XX (UABR-) UAB RES FOUND.
 PA (UVFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Prognuske-Fox A;
 PT Tumasorn S;
 XX WPI: 1996-287181/29.
 DR N-PSDB: AAT30648.
 CC Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS
 XX Claim 4: Page 114-115; 153pp; English.
 CC HAREP4 (AAR96024) is the product of the HAREP4 repeat unit (AAT30648)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP4 and other
 CC hga repeat unit products (see also AAR96021-23) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of live vaccine. HAREP4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 CC
 SQ Sequence 439 AA:
 OY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 Query Match 100.0%; Score 78; DB 17; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 5
 ID AAR96492 standard; Protein: 439 AA.
 XX AAR96492;
 AC
 XX 22-DEC-1998 (first entry)
 DT
 XX Haemagglutinin protein hga, HAREP4.
 DE
 XX Haemagglutinin protein; periodontal disease; vaccine; hga.
 KM
 XX Porphyromonas gingivalis.
 OS
 XX US5824791-A.
 PM
 XX 20-OCT-1998.
 PD
 XX 11-DEC-1995: 95US-0570311.
 PF
 XX 11-DEC-1995: 95US-0570311.
 PR 08-SEP-1988: 88US-0241640.
 PR 25-JAN-1991: 91US-0647119.
 PR 09-DEC-1994: 94US-0353485.
 XX (UABR-) UAB RES FOUND.

PA (UVFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Prognuske-Fox A;
 PT Tumasorn S;
 XX WPI: 1996-287181/29.
 DR N-PSDB: AAT30648.
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hga haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 CC
 SQ Sequence 439 AA:
 OY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 Query Match 100.0%; Score 78; DB 19; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 6
 ID AAR96021 standard; Protein: 450 AA.
 XX AAR96021;
 AC
 XX 04-SEP-1996 (first entry)
 DT
 XX P. gingivalis haemagglutinin hga HAREP1 product.
 DE
 XX Haemagglutinin: hga; periodontal disease; vaccine; antibody;
 KM HAREP1.
 XX Porphyromonas gingivalis strain 381.
 OS
 XX MO9617936-A2.
 PM
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995: 95WO-US16108.
 PF
 XX 09-DEC-1994: 94US-0353485.
 PR
 XX (UABR-) UAB RES FOUND.
 PA (UVFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Prognuske-Fox A;
 PT Tumasorn S;
 XX WPI: 1996-287181/29.
 DR N-PSDB: AAT30645.
 CC Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS
 XX Claim 4: Page 103-104; 153pp; English.
 CC HAREP1 (AAR96021) is the product of the HAREP1 repeat unit (AAT30645)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP1 and other
 CC hga repeat unit products (see also AAR96022-24) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or

CC animals against periodontal disease. Expression in *Salmonella*
 CC cells allows production of live vaccine. Harepi-4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic apn.

SO Sequence 450 AA:

Query Match 100.0%: Score 78: DB 17: Length 450:

Best Local Similarity 100.0%: Pred. No. 8.2e-06: Mismatches 0: Gaps 0:

Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ALPNPNYLISKDVTG 15
 DB 245 ALPNPNYLISKDVTG 259

RESULT 7

AAW69489 standard: Protein: 450 AA.

AC AAW69489:

DT 22-DEC-1998 (first entry)

DE Haemagglutinin protein hagg, Harepi.

KW Haemagglutinin protein: periodontal disease; vaccine; hagg.

OS Porphyromonas gingivalis.

PN US824791-A.

PD 20-OCT-1998.

PF 11-DEC-1995: 95US-0570311.

PR 11-DEC-1995: 95US-0570311.

PR 08-SEP-1988: 88US-0241640.

PR 25-JAN-1991: 91US-0647119.

PR 09-DEC-1994: 94US-0353485.

PA (UABR-) UAB RES FOUND.

PA (UYFL) UNIV FLORIDA.

PI Han N. Lantz M, Lepine G, Patti JM, Proguiske-Fox A;

PI Tummasorn S;

PS WPI: 1998-582627/49.

DR N-PSDB: AAV58876.

XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin

XX and/or protease poly(peptide(s))

PS Claim 1: Column 121-126: 101pp: English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the

XX invention. This sequence represents the hagg haemagglutinin protein. The

XX polypeptides are used to produce antibodies to organisms associated with

XX periodontal disease. The antibodies are also used in purification and

XX identification procedures. The genes and polypeptides are used as

XX vaccines against periodontal disease.

SO Sequence 450 AA:

Query Match 100.0%: Score 78: DB 19: Length 450:

Best Local Similarity 100.0%: Pred. No. 8.2e-06: Mismatches 0: Gaps 0:

Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ALPNPNYLISKDVTG 15
 DB 245 ALPNPNYLISKDVTG 259

RESULT 8
 AAR96022 standard: Protein: 456 AA.

AC AAR96022:

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin hagg Harep2 product.

KW Haemagglutinin: hagg, periodontal disease; vaccine; antibody;

OS Porphyromonas gingivalis strain 381.

PN WO9617936-A2.

PD 13-JUN-1996.

PF 11-DEC-1995: 95WO-US16108.

PR 09-DEC-1994: 94US-0353485.

PA (UABR-) UAB RES FOUND.

PA (UYFL) UNIV FLORIDA.

PI Han N. Lantz M, Lepine G, Patti JM, Proguiske-Fox A;

PI Tummasorn S;

DR WPI: 1996-287181/29.

DR N-PSDB: AAT30646.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4: Page 107-108: 153pp: English.

XX Harep2 (AAR96022) is the product of the Harep2 repeat unit (AAT30646)

XX of the hagg gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin hagg (see also AAR96030). Harep2 and other hagg

XX repeat unit products (see also AAR96021 and AAR96023-24) can be obd.

XX from transformed host cells and used as vaccines to protect humans

XX or animals against periodontal disease. Expression in *Salmonella*

XX cells allows production of live vaccine. Harepi-4 can also be used

XX to detect the presence of anti-P. gingivalis antibodies and to

XX raise monoclonal antibodies for diagnostic apn.

SO Sequence 456 AA:

Query Match 100.0%: Score 78: DB 17: Length 456:

Best Local Similarity 100.0%: Pred. No. 8.3e-06: Mismatches 0: Gaps 0:

Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ALPNPNYLISKDVTG 15
 DB 251 ALPNPNYLISKDVTG 265

PS Claim 9

XX AAR96023 standard: Protein: 456 AA.

XX AAR96023:

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin hagg Harep3 product.

KW Haemagglutinin: hagg, periodontal disease; vaccine; antibody;

OS Porphyromonas gingivalis strain 381.

PN M09617936-A2.
 PD 13-JUN-1996.
 XX 11-DEC-1995; 95WO-US16108.
 XX 09-DEC-1994; 94US-0353485.
 XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1996-287181/29.
 DR N-PSDB; AAT30647.
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 110-112; 153pp; English.
 XX HAREP3 (AAR96023) is the product of the HAREP3 repeat unit (AAT30647)
 CC of the haggA gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin haggA (see also AAR96030). HAREP3 and other haggA
 CC repeat unit products (see also AAR96021-22 and AAR96024) can be obtd.
 CC from transformed host cells and used as vaccines to protect humans
 CC or animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of live vaccine. HAREP1-4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 SO Sequence 456 AA;
 Query Match 100.0%; Score 78; DB 17; Length 456;
 Best Local Similarity 100.0%; Pred. No. 8.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 10
 AAM69490 standard; Protein: 456 AA.
 XX AAM69490;
 AC 22-DEC-1998 (first entry)
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein haggA, HAREP2.
 XX Haemagglutinin protein; periodontal disease; vaccine; haggA.
 KM Haemagglutinin protein; periodontal disease; vaccine; haggA.
 XX Porphyromonas gingivalis.
 OS US5824791-A.
 PM 20-OCT-1998.
 PD 11-DEC-1995; 95US-0570311.
 PF 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;

XX WPI: 1998-582627/49.
 DR N-PSDB; AAV58877.
 XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s))
 XX Claim 1; Column 127-132; 101pp; English.
 XX This sequence is encoded by a porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haggA haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 SO Sequence 456-AA;
 Query Match 100.0%; Score 78; DB 19; Length 456;
 Best Local Similarity 100.0%; Pred. No. 8.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 11
 AAM69491 standard; Protein: 456 AA.
 XX AAM69491;
 AC 22-DEC-1998 (first entry)
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein haggA, HAREP3.
 XX Haemagglutinin protein; periodontal disease; vaccine; haggA.
 KM Haemagglutinin protein; periodontal disease; vaccine; haggA.
 XX Porphyromonas gingivalis.
 OS US5824791-A.
 PM 20-OCT-1998.
 PD 11-DEC-1995; 95US-0570311.
 PF 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB; AAV58878.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s))
 PS Claim 1; Column 133-138; 101pp; English.
 XX This sequence is encoded by a porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haggA haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

SO Sequence 456 AA:

Query Match 100.0%; Score 78; DB 19; Length 456;
 Best Local Similarity 100.0%; Pred. No. 8.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTG 15
 |||
 DB 251 ALNPDNYLISKDVTG 265

RESULT 12
 AAR96028
 ID AAR96028 standard; Protein: 1087 AA.
 AC AAR96028:
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haemagglutinin haGD.
 KW Haemagglutinin; haGD; periodontal disease; vaccine; antibody.
 OS Porphyromonas gingivalis strain FDC381.
 PN MO9617936-A2.
 PD 13-JUN-1996.
 PF 11-DEC-1995; 95WO-US16108.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patil JM, Proguiske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1996-287181/29.
 DR N-PSDB: AAT30652.
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 5; Page 65-68; 153pp; English.
 CC P. gingivalis 381 haemagglutinin haGD (AAR96028) was identified as
 CC the product of a gene (AAT30652) isolated from a P. gingivalis 318
 CC genomic library. The haemagglutinin (see also AAR96032) can be obt.
 CC from transformed host cells and used as a vaccine to protect humans
 CC or animals against periodontal disease. Expression in Salmonella
 CC cells allows prodn. of a live vaccine. The haemagglutinin can also
 CC be used to detect the presence of anti-P. gingivalis antibodies and
 CC to raise monoclonal antibodies for diagnostic appln.
 SQ Sequence 1087 AA:

Query Match 100.0%; Score 78; DB 17; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTG 15
 |||
 DB 579 ALNPDNYLISKDVTG 593

RESULT 13
 AAM69486
 ID AAM69486 standard; Protein: 1087 AA.
 AC AAM69486:
 XX

DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein haGD.
 KW Haemagglutinin protein; periodontal disease; vaccine; haGD.
 OS Porphyromonas gingivalis.
 PN US5824791-A.
 PD 20-OCT-1998.
 PF 11-DEC-1995; 95US-0570311.
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patil JM, Proguiske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB: AAV58873.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))
 PS Claim 1; Column 57-64; 101pp; English.
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haGD haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 SQ Sequence 1087 AA:

Query Match 100.0%; Score 78; DB 19; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTG 15
 |||
 DB 579 ALNPDNYLISKDVTG 593

RESULT 14
 AAR96032
 ID AAR96032 standard; Protein: 1358 AA.
 AC AAR96032:
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haGD haemagglutinin.
 KW Haemagglutinin; haGD; periodontal disease; vaccine; antibody.
 OS Porphyromonas gingivalis strain FDC381.
 PN MO9617936-A2.
 PD 13-JUN-1996.
 PF 11-DEC-1995; 95WO-US16108.
 PR 09-DEC-1994; 94US-0353485.
 XX

PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patil JM, Proguliske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1996-287181/29.
 DR N-PSDB: AAT30655.
 XX
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 and vaccination against periodontal disease
 XX
 PS Claim 5: Page 125-129; 153pp; English.
 CC P. gingivalis 381 haemagglutinin hagd (AAR96032) was identified as
 CC the product of the second open reading frame of the hagd gene
 CC (AAT30655) derived from P. gingivalis 318 genomic DNA. A first
 CC open reading frame coded for hagd protease (see also AAR96031).
 CC The protease and haemagglutinin can be obtd. from transformed host
 CC cells and used in vaccines to protect humans or animals against
 CC periodontal disease. Expression in Salmonella cells allows produ.
 CC of live vaccines. The haemagglutinin and protease can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic appln.
 XX
 SQ Sequence 1358 AA;

Query Match 100.0%; Score 78; DB 17; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 3.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 850 ALNPDNYLISKDVTG 864

RESULT 15

AAM69494
 ID AAM69494 standard; Protein: 1358 AA.

AC AAM69494;

DT 22-DEC-1998 (first entry)

DE Haemagglutinin protein hagd.

KW Haemagglutinin protein; periodontal disease; vaccine; hagd.

OS Porphyromonas gingivalis.

PN US5824791-A.

PD 20-OCT-1998.

PF 11-DEC-1995; 95US-0570311.

PR 11-DEC-1995; 95US-0570311.

PR 08-SEP-1988; 88US-0241640.

PR 25-JAN-1991; 91US-0647119.

PR 09-DEC-1994; 94US-0353485.

PA (UABR-) UAB RES FOUND.

PA (UYFL) UNIV FLORIDA.

PI Han N, Lantz M, Lepine G, Patil JM, Proguliske-Fox A;

PI Tumwasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB: AAV58880.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 and/or protease poly(peptide(s))

PS Claim 1: Column 145-158; 101pp; English.

CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hagd haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX

SQ Sequence 1358 AA;

Query Match 100.0%; Score 78; DB 19; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 3.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 850 ALNPDNYLISKDVTG 864

Search completed: June 26, 2003, 00:51:42
 Job time: 6.29412 secs

Thu Jun 26 11:58:03 2003

us-09-980-370-1.rapp

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:39 ; Search time 3.17647 Seconds
(without alignments)
510.977 Million cell updates/sec

Title: US-09-980-370-1
Sequence: 1 ALPNDYLSKDVTK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCYT_NEM_PUB pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCU05_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	53.8	2548	10 US-09-851-682A-1	Sequence 1, Appl1
2	41	52.6	166	9 US-09-925-299-1177	Sequence 1177, Ap
3	41	52.6	166	10 US-09-925-299-1177	Sequence 1177, Ap
4	40	51.3	575	9 US-09-738-626-4263	Sequence 4263, Ap
5	40	51.3	1725	9 US-10-037-182-12	Sequence 12, Appl
6	40	51.3	1765	9 US-10-037-182-8	Sequence 8, Appl
7	40	51.3	1766	9 US-10-037-182-6	Sequence 6, Appl
8	40	51.3	1766	9 US-10-037-182-10	Sequence 10, Appl
9	40	51.3	1766	9 US-09-873-676-113	Sequence 113, Appl
10	40	51.3	1766	10 US-09-938-275-6	Sequence 6, Appl1
11	40	51.3	1766	10 US-09-938-275-7	Sequence 7, Appl1
12	39	50.0	72	10 US-09-205-658-178	Sequence 178, App
13	39	50.0	90	10 US-09-205-658-191	Sequence 191, App
14	39	50.0	331	9 US-10-051-643-182	Sequence 182, App
15	39	50.0	331	9 US-09-880-505-182	Sequence 182, App
16	39	50.0	368	10 US-09-801-368-308	Sequence 308, App
17	39	50.0	632	10 US-09-205-658-159	Sequence 159, App
18	39	50.0	636	10 US-09-205-658-160	Sequence 160, App
19	38	48.7	93	10 US-09-731-872-452	Sequence 452, App

20	38	48.7	381	10 US-09-815-242-10962	Sequence 10962, A
21	38	48.7	613	10 US-09-862-027-39	Sequence 39, Appl
22	38	48.7	1694	9 US-10-223-070-19	Sequence 19, Appl
23	37	47.4	355	10 US-09-803-286A-12	Sequence 12, Appl
24	37	47.4	357	10 US-09-947-027-9	Sequence 9, Appl1
25	37	47.4	357	12 US-10-091-009-9	Sequence 9, Appl1
26	37	47.4	822	10 US-09-740-627-11	Sequence 11, Appl
27	37	47.4	822	9 US-10-081-119-16	Sequence 16, Appl
28	36.5	46.8	109	10 US-09-867-550-474	Sequence 474, App
29	36.5	46.8	491	10 US-09-881-752A-128	Sequence 128, App
30	36	46.2	118	10 US-10-106-698-6332	Sequence 6332, Ap
31	36	46.2	120	10 US-09-764-903-62	Sequence 62, App
32	36	46.2	135	9 US-09-764-868-830	Sequence 830, App
33	36	46.2	139	9 US-09-901-938-14	Sequence 14, App
34	36	46.2	160	10 US-09-764-877-1284	Sequence 1284, Ap
35	36	46.2	168	10 US-10-081-347-22	Sequence 22, Appl
36	36	46.2	168	10 US-09-750-963-13	Sequence 13, Appl
37	36	46.2	181	10 US-09-902-773A-2	Sequence 2, Appl1
38	36	46.2	181	10 US-09-425-021-18	Sequence 18, Appl
39	36	46.2	207	10 US-09-778-927A-65	Sequence 65, Appl
40	36	46.2	243	9 US-10-081-347-21	Sequence 21, Appl
41	36	46.2	243	10 US-09-822-485-15	Sequence 15, Appl
42	36	46.2	243	10 US-09-251-263-2	Sequence 2, Appl1
43	36	46.2	298	10 US-09-801-368-66	Sequence 66, Appl
44	36	46.2	332	9 US-09-984-130-32	Sequence 32, Appl
45	36	46.2	332	9 US-10-215-457-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-851-682A-1
Sequence 1, Application US/09851682A
Patient No. US2002091248A1
GENERAL INFORMATION:
APPLICANT: Adams, Arwen E.
APPLICANT: Chiu, Choi Ying
APPLICANT: Duhl, David
APPLICANT: Gorman, Susan W.
APPLICANT: Leng, Song
APPLICANT: Shelton, Val
APPLICANT: Welch, Juliet
TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
FILE REFERENCE: 200130.442
CURRENT APPLICATION NUMBER: US/09/851.682A
CURRENT FILING DATE: 2001-05-08
PRIOR APPLICATION NUMBER: US/09/172.422
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2548
TYPE: PRT
ORGANISM: Homo sapien
US-09-851-682A-1
Query Match 53.8%; Score 42; DB 10; Length 2548;
Best Local Similarity 70.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 2 LNPDNYLSK 11
DB 989 LNPDNYGVK 998
RESULT 2
US-09-925-299-1177
Sequence 1177, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:

Thu Jun 26 11:58:03 2003

us-09-980-370-1.rapb

Page 2

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1177
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (157)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (158)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (163)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; US-09-925-299-1177

Query Match          52.6%: Score 41; DB 9; Length 166;
Best Local Similarity 61.5%: Pred. No. 16;
Matches      8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 ALMPDNYLSKDV 13

```

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1177
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (157)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (158)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (163)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; US-09-925-299-1177

Query Match          52.6%: Score 41; DB 9; Length 166;
Best Local Similarity 61.5%: Pred. No. 16;
Matches      8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 ALMPDNYLSKDV 13

```

OY 2 LNPBNYLISKDVT 14
||||:|
DB 28 LNPDSHLIENVVT 40

RESULT 6

US-10-037-182-6
; Sequence 8, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Tryggevason, Karl
; APPLICANT: Thyboll, Jili
; TITLE OF INVENTION: Recombinant Lamnin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-037-182-6

Query Match 51.3%; Score 40; DB 9; Length 1765;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LNPBNYLISKDVT 14
||||:|
DB 68 LNPDSHLIENVVT 80

RESULT 7

US-10-037-182-6
; Sequence 6, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Tryggevason, Karl
; APPLICANT: Thyboll, Jili
; TITLE OF INVENTION: Recombinant Lamnin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-037-182-6

Query Match 51.3%; Score 40; DB 9; Length 1786;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LNPBNYLISKDVT 14
||||:|
DB 89 LNPDSHLIENVVT 101

RESULT 8

US-10-037-182-10

; Sequence 10, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Tryggevason, Karl
; APPLICANT: Thyboll, Jili
; TITLE OF INVENTION: Recombinant Lamnin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/279,282
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-037-182-10

Query Match 51.3%; Score 40; DB 9; Length 1786;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LNPBNYLISKDVT 14
||||:|
DB 89 LNPDSHLIENVVT 101

RESULT 9

US-09-873-676-113
; Sequence 113, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 113
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-676-113

Query Match 51.3%; Score 40; DB 10; Length 1786;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LNPBNYLISKDVT 14
||||:|
DB 89 LNPDSHLIENVVT 101

RESULT 10

US-09-938-275-6
; Sequence 6, Application US/09938275
; Patent No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Castillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; TITLE OF INVENTION: of laminin and laminin-derived Protein Fragments

Thu Jun 26 11:58:03 2003

us-09-980-370-1.rapb

Page 4

```
FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938.275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1786
TYPE: PRT
ORGANISM: Homo Sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER:
DATABASE ENTRY DATE: 1988-08-01
US-09-938-275-6
```

```
Query Match
Best Local Similarity 51.3%; Score 40; DB 10; Length 1786;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 LNPBNYLISKDVT 14
DB 89 LNPDSHLLENVVT 101
```

```
RESULT 11
US-09-938-275-7
Sequence 7, Application US/09938275
Patent No. US20020111309A1
GENERAL INFORMATION:
APPLICANT: Gerardo Castillo
TITLE OF INVENTION: Therapeutic and Diagnostic Applications
FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938.275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1786
TYPE: PRT
ORGANISM: Mus Musculus
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Swissprot P02469
DATABASE ENTRY DATE: 1989-07-01
US-09-938-275-7
```

```
Query Match
Best Local Similarity 51.3%; Score 40; DB 10; Length 1786;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 LNPBNYLISKDVT 14
DB 89 LNPDSHLLENVVT 101
```

```
RESULT 12
US-09-205-658-178
Sequence 178, Application US/09205658
Patent No. US20010029617A1
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351004
CURRENT APPLICATION NUMBER: US/09/205.658
CURRENT FILING DATE: 1998-12-03
EARLIER APPLICATION NUMBER: 08/857,076
EARLIER FILING DATE: 1997-05-15
EARLIER APPLICATION NUMBER: 08/888,534
EARLIER FILING DATE: 1997-07-07
EARLIER APPLICATION NUMBER: US98/10080
EARLIER FILING DATE: 1998-05-15
```

```
NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 178
LENGTH: 72
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-205-658-178
```

```
Query Match
Best Local Similarity 50.0%; Score 39; DB 10; Length 72;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 LNPBNYLISKD 12
DB 50 MKPDNVLIOKD 60
```

```
RESULT 13
US-09-205-658-191
Sequence 191, Application US/09205658
Patent No. US20010029617A1
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351004
CURRENT APPLICATION NUMBER: US/09/205.658
CURRENT FILING DATE: 1998-12-03
EARLIER APPLICATION NUMBER: 08/857,076
EARLIER FILING DATE: 1997-05-15
EARLIER APPLICATION NUMBER: 08/888,534
EARLIER FILING DATE: 1997-07-07
EARLIER APPLICATION NUMBER: US98/10080
EARLIER FILING DATE: 1998-05-15
NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 191
LENGTH: 90
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-205-658-191
```

```
Query Match
Best Local Similarity 50.0%; Score 39; DB 10; Length 90;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 LNPBNYLISKD 12
DB 68 MKPDNVLIOKD 78
```

```
RESULT 14
US-10-051-643-182
Sequence 182, Application US/10051643
Publication No. US20020197265A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Methods and Compounds for the Treatment
TITLE OF INVENTION: of Immunologically-mediated Diseases of the Respiratory
FILE REFERENCE: 11000.10082
CURRENT APPLICATION NUMBER: US/10/051.643
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US09/156,181
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 08/996,624
PRIOR FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 182
LENGTH: 331
```


TYPE: PRT
ORGANISM: Mycobacterium vaccae
US-10-051-643-182

Query Match 50.0%; Score 39; DB 9; Length 331;
Best Local Similarity 63.6%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 PDNYLISKDVT 14
||:|:|:|
Db 244 PDHYLVLSLVT 254

RESULT 15
US-09-880-505-182
Sequence 182, Application US/09880505
Publication No. US20030007976A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L.J.
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Methods and Compounds for the Treatment
FILE REFERENCE: 11000.1007c2
CURRENT APPLICATION NUMBER: US/09/880,505
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 09/324,542
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: US 08/997,080
PRIOR FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 194
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 182
LENGTH: 331
TYPE: PRT
ORGANISM: Mycobacterium vaccae
US-09-880-505-182

Query Match 50.0%; Score 39; DB 9; Length 331;
Best Local Similarity 63.6%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 PDNYLISKDVT 14
||:|:|:|
Db 244 PDHYLVLSLVT 254

Search completed: June 26, 2003, 00:54:44
Job time : 3.17647 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:46:24 ; Search time 2.47059 seconds

(Without alignments)
583.673 Million cell updates/sec

Title: US-09-980-370-1

Perfect score: 78
Sequence: 1 ALNPDNYLISKDVTC 15

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	78	100.0	1526 2	gingipain R (EC 3.4.22.37)
2	78	100.0	1704 2	gingipain R (EC 3.4.22.37)
3	78	100.0	1732 2	lysine-specific cy
4	78	100.0	2628 2	hemagglutinin A -
5	47	60.3	334 2	hypothetical prote
6	45	57.7	265 2	tubulin gamma chain
7	45	57.7	315 2	clonamide alcohol d
8	44	56.4	315 2	hypothetical prote
9	44	56.4	639 2	phosphoenolpyruvat
10	44	56.4	752 2	kinase-like protei
11	43	55.1	280 2	protein kinase hom
12	43	55.1	280 2	hypothetical prote
13	42	53.8	318 2	hypothetical prote
14	42	53.8	561 2	hypothetical prote
15	42	53.8	1777 2	hypothetical prote
16	42	53.8	2548 2	myosin IXA (import
17	41	52.6	146 2	protein kinase hom
18	41	52.6	314 2	growth arrest-spec
19	41	52.6	363 2	hypothetical prote
20	41	52.6	393 2	probable protein k
21	41	52.6	821 2	phosphoenolpyruvat
22	41	52.6	829 2	periplasmic nitrat
23	41	52.6	834 2	Napa periplasmic n
24	40	51.3	883 2	hypothetical prote
25	40	51.3	1786 1	laminin beta-1 cha
26	40	51.3	1786 1	laminin beta-1 cha
27	39.5	50.6	2626 2	myosin-RhGAP prot
28	39.5	50.6	654 2	phosphoglycerate k
29	39.5	50.6	189 2	hypothetical prote

30	39	50.0	274 2	S28039	nuclease nuca prec
31	39	50.0	274 2	A82523	sugar-non-specific
32	39	50.0	312 2	T32446	hypothetical prote
33	39	50.0	312 2	A89460	protein H4X12.1 f
34	39	50.0	321 2	T24599	hypothetical prote
35	39	50.0	326 2	B46108	outer capsid prote
36	39	50.0	329 2	T46214	hypothetical prote
37	39	50.0	368 2	A54430	hypoxic function t
38	39	50.0	372 1	A54870	cystathionine beta
39	39	50.0	472 1	A54870	cell adhesion rece
40	39	50.0	636 2	G81983	pilin glycosylatio
41	39	50.0	694 2	E69143	hypothetical prote
42	38.5	49.4	13055 2	T16580	hypothetical prote
43	38.5	49.4	302 2	T11675	lactoyglutathione
44	38.5	49.4	917 2	B81309	isoenzyme-cRNA 11
45	38	48.7	181 2	T02591	germin-like protei

ALIGNMENTS

RESULT 1

549763
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)

C:Species: Porphyromonas gingivalis
C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997

C:Accession: 549763
R:Aduse-Opoku, J.; Mull, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.

Submitted to the EMBL Data Library, November 1994
A:Description: Cloning, sequence analysis and expression in Escherichia coli of pprp1

A:Reference number: 549763

A:Accession: 549763

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1526 <ADU>

A:Cross-references: EMBL:X82680

C:Genetics:

A:Gene: pprp1

C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0% Score 78; DB 2; Length 1526;

Best Local Similarity 100.0%; Pred. No. 8; 8e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1198 ALNPDNYLISKDVTC 1212

QY 1 ALNPDNYLISKDVTC 15

DB 1198 ALNPDNYLISKDVTC 1212

RESULT 2

A55426
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis

N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase: HGP

C:Species: Porphyromonas gingivalis
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999

C:Accession: A55426; D53113
R:Payloft, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr

J. Biol. Chem. 270, 1007-1010, 1995
A:Title: Molecular cloning and structural characterization of the Arg-gingipain prote

A:Reference number: A55426; MIMD:95138080; PMID:7836351

A:Accession: A55426

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1704 <PAV>

A:Cross-references: GB:015282; NID:9557067; PIDN:AAA69539.1; PID:9557068

J. Pike, R.; McGraw, W.; Potempa, J.; Travis, J.

J. Biol. Chem. 269, 406-411, 1994

A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso

A:Reference number: A53113; MIMD:94103245; PMID:8276827

A:Accession: D53113

A:Status: preliminary

A:Molecule type: protein

A:Residues: 228-249 <PIK>

A: Experimental source: H66
A: Note: sequence extracted from NCBI backbone (NCBI:U141694)
C: Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 78; DB 2; Length 1704;
Best Local Similarity 100.0%; Pred. No. 9.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPNPNYLISKDVTG 15
DB 1196 ALPNPNYLISKDVTG 1210

RESULT 3

T30816
Lysine-specific cysteine proteinase porphyrin (EC 3.4.22.-) - Porphyromonas gingivalis
N: Alternate names: lysine-specific cysteine proteinase 1, 60k
C: Species: Porphyromonas gingivalis
C: Date: 22-Oct-1999 #sequence-revision 22-Oct-1999 #text-change 17-Nov-2000

A: Accession: T30836; #20837; 130526; AS3113
R: Barcoy-Gallagher, G.A.; Han, J.M.; Whitlock, J.; Proguiske-Fox, A.; Lantz, J.
A: Title: Analysis of the prtP gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
A: Reference number: 220893; MUID: 96213011; PMID: 8631659

A: Accession: T30836
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1732 <GB>
A: Cross-references: EMBL:U42210; NID: g1314325; PID: g1314326; PIDN: AAB06565.1

R: Stakel, N.; Chai, S.M.; Reynolds, E.C.
A: Submitted to the EMBL Data Library, October 1996
A: Reference number: 220896
A: Accession: T30937
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-795; 1-797-1389; 'N', 1391-1478; 'Y', 1480-1732 <SLA>

A: Cross-references: EMBL:U75366; NID: g2182812; PID: AAB0809.1
R: Levy, J.P.; Macrina, G.L.
A: Title: Immun 68, 3035-3042, 1996
A: Reference number: 220844; MUID: 9828016; PMID: 9632363

A: Accession: T30926
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEW>

A: Cross-references: EMBL:AF017059; NID: g2738802; PID: g2738803; PIDN: AAC26523.1
R: Eke, R.; McGraw, W.; Peltz, J.; Travis, J.
A: Title: Chem 269, 406-411, 1994

A: Reference number: AS3113; MUID: 94103245; PMID: 8276827
A: Accession: AS3113
A: Status: preliminary
A: Molecule type: protein
A: Residues: 229-249 <PTK>
A: Experimental source: H66
A: Note: sequence extracted from NCBI backbone (NCBI:U141690)

C: Keywords: cysteine proteinase; hydrolase
A: Gene: prtP
Query Match 100.0%; Score 78; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPNPNYLISKDVTG 15
DB 1216 ALPNPNYLISKDVTG 1230

RESULT 4

T28651
hemagglutinin A - Porphyromonas gingivalis
C: Species: Porphyromonas gingivalis

C: Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 03-Aug-2001
A: Accession: T28651
R: Han, N.; Whitlock, J.; Proguiske-Fox, A.

A: Title: Immun 64, 4000-4007, 1996
A: Title: The hemagglutinin gene A (haga) of Porphyromonas gingivalis 361 contains four
A: Reference number: 220494; MUID: 97047672; PMID: 8928061

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>

A: Cross-references: EMBL:U41807; NID: g1552410; PID: g1469916; PIDN: AAB17128.1
C: Genes: haga

Query Match 100.0%; Score 78; DB 2; Length 2628;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPNPNYLISKDVTG 15
DB 744 ALPNPNYLISKDVTG 758

RESULT 5

T27658
hypothetical protein ZK1037.3 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 15-Oct-1999

R: Basham, V.
A: Submitted to the EMBL Data Library, October 1996
A: Reference number: 220401
A: Accession: T27658
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-334 <WIL>

A: Cross-references: EMBL:Z81142; PIDN: CAB03502.1; GSPDB: GN00023; CESP: ZK1037.3
A: Experimental source: clone ZK1037
C: Genes: ZK1037
A: Gene: CESP: ZK1037.3
A: Map position: 5

A: Introns: 57/3; 199/3; 248/3; 292/1
Query Match 60.3%; Score 47; DB 2; Length 334;
Best Local Similarity 69.2%; Pred. No. 3.2;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LNPDPYLCSDVTG 14
DB 22 LNPDPYLCSDVTG 34

RESULT 6

S40209
tubulin gamma chain - fungus (Cochliobolus heterostrophus)
C: Species: Cochliobolus heterostrophus; Bipolaris maydis
C: Date: 06-Jan-1995 #sequence-revision 06-Jan-1995 #text-change 13-Aug-1999

A: Accession: S40209
R: Parkinson, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
A: Submitted to the EMBL Data Library, August 1993
A: Description: Phylogenetic analyses using the gamma tubulin gene.

A: Reference number: S40209
A: Accession: S40209
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-265 <PAR>

A: Cross-references: EMBL: X74455; NID: g437988; PIDN: CAA52464.1; PID: g437989
C: Genes: tubulin
A: Introns: 136/3
C: Superfamily: tubulin

Query Match 57.7%; Score 45; DB 2; Length 265;
Best Local Similarity 61.5%; Pred. No. 5.5;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 NPDNYLSKQVTG 15
 |||:|||||
 Db 22 NPNFYTHKQGTG 34

RESULT 7

cinamyl-alcohol dehydrogenase (EC 1.1.1.195) - alfalfa

C:Species: Medicago sativa (alfalfa)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 15-Oct-1999
 C:Accession: S31572
 R:van Doorslaere, J., van Montagu, M., Inze, D.
 submitted to the EMBL Data Library, January 1993
 A:Description: isolation and characterization of poplar and alfalfa cinamyl alcohol dehydrogenase
 A:Reference number: S31571
 A:Accession: S31572
 A:Molecule type: mRNA
 A:Residues: 1-358 <VNA>
 A:Cross-references: EMBL:219573; NID:919594; PIDN:CAA79625.1; PID:919595
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 C:Keywords: alcohol metabolism; metalloprotein; oxidoreductase; zinc
 F:33-34/Domains: long-chain alcohol dehydrogenase homology <LADH>
 F:48-70/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted

Query Match 57.7%; Score 45; DB 2; Length 358;
 Best Local Similarity 57.1%; Pred. No. 7.7;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 LMPDNYLSKQVTG 15
 |||:|||||
 Db 225 LGADNYLVSSDTYG 238

RESULT 8

hypothetical protein F14J22.12 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96533
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; WUID:21016719; PMID:11130712
 A:Accession: B96533
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-315 <STO>
 A:Cross-references: GB:AE005173; NID:910120425; PIDN:AA613050.1; GSPDB:GN00141
 C:Gene: F14J22.12
 A:Map position: 1

Query Match 56.4%; Score 44; DB 2; Length 315;
 Best Local Similarity 61.5%; Pred. No. 9.8;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ALPNDNYLSKQVD 13
 |||:|||||
 Db 34 SLNPNDNYVSKDV 46

RESULT 9

hypothetical protein yscA [Imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: A86845
 R:Botolin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissbach, J.; Eh Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
 A:Reference number: A86845; WUID:2125186; PMID:11337471
 A:Accession: A86845
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-639 <STO>
 A:Cross-references: GB:AE005176; PID:912724783; PIDN:AAK05859.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Gene: yscA

Query Match 56.4%; Score 44; DB 2; Length 639;
 Best Local Similarity 81.8%; Pred. No. 21;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 PDNYLSKQVT 14
 |||:|||||
 Db 452 PDNYLSKQVT 462

RESULT 10

phosphoenolpyruvate synthase [Imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: F84192
 R:Ng, M.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Lettner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; A:Reference number: A84160; WUID:20504483; PMID:11016950
 A:Accession: F84192
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-752 <STO>
 A:Cross-references: GB:AE004437; NID:910579956; PIDN:AA618906.1; GSPDB:GN00138
 C:Gene: ppsA
 C:Superfamily: Escherichia coli pyruvate, water dikinase; phosphotransferase system en

Query Match 56.4%; Score 44; DB 2; Length 752;
 Best Local Similarity 50.0%; Pred. No. 26;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 ALPNDNYLSKD 12
 |||:|||||
 Db 217 AVSPDNTVDRD 228

RESULT 11

kinase-like protein [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
 C:Accession: H85215
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; WUID:20083488; PMID:10617198
 A:Accession: H85215
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <STO>
 A:Cross-references: GB:NC_001268; NID:97268706; PIDN:CAB78913.1; GSPDB:GN00140
 C:Gene: Atg19110

A:Map position: 4
C:Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match
Best Local Similarity 55.1%; Score 43; DB 2; Length 290;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMPDNYLISKDV 13
| | | | | | | | | |
DB 126 LKPNLVLVSKDI 137

RESULT 12

T04432
protein kinase homolog T18B16.80 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: T04432
R:Bevan, M.; Benes, V.; Rechmann, S.; Borrova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.

submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15355

A:Accession: T04432
A:Molecule type: DNA

A:Residues: 1-290 <BEV>

A:Cross-references: EMBL:AL021687

A:Experimental source: cultivar Columbia; BAC clone T18B16

C:Genetics:
A:Map position: 4
A:Note: T18B16.80; Intron positions not resolved

Query Match
Best Local Similarity 55.1%; Score 43; DB 2; Length 290;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMPDNYLISKDV 13
| | | | | | | | | |
DB 126 LKPNLVLVSKDI 137

RESULT 13

hypothetical protein T29M8.6 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: D86325
R:Phelios, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, N.F.; Hughes, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
Nansen, C.W.; Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

A:Authors: Salzman, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.-Y.; G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; M01D:21016719; PMID:11130712

A:Accession: D86325
A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-318 <STO>

A:Cross-references: GB:AE005172; NID:G8954057; PIDN:AMF82230.1; GSPDB:GN00141

C:Genetics:
A:Map position: 1

Query Match
Best Local Similarity 53.8%; Score 42; DB 2; Length 318;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPNYLYISKD 12
: | | | | | | | | | |
DB 33 SLNPNGVYVSKD 44

RESULT 14

T34368
hypothetical protein T19D12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34368
R:Favell, A.
submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid T19D12.
A:Reference number: Z21513

A:Accession: T34368
A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-561 <FAV>

A:Cross-references: EMBL:U41263; PIDN:AA24423.1; GSPDB:GN00020; CESP:T19D12.2

A:Experimental source: strain Bristol N2; clone T19D12

C:Genetics:
A:Gene: CESP:T19D12.2
A:Map position: 2
A:Introns: 25/1; 151/2; 214/1; 282/3; 333/2; 443/2; 469/2; 492/2

Query Match
Best Local Similarity 53.8%; Score 42; DB 2; Length 561;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 NPDPYLYISKDV 15
| | | | | | | | | |
DB 295 NPDSYGLNDVNG 307

RESULT 15

T34369
hypothetical protein T19D12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34369
R:Favell, A.

submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid T19D12.

A:Reference number: Z21513

A:Accession: T34369
A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-1777 <FAV>

A:Cross-references: EMBL:U41263; PIDN:AA24428.1; GSPDB:GN00020; CESP:T19D12.1

A:Experimental source: strain Bristol N2; clone T19D12

C:Genetics:
A:Gene: CESP:T19D12.1
A:Map position: 2

A:Introns: 36/1; 134/2; 180/1; 622/3; 691/2; 754/1; 1111/2; 1174/1; 1271/3; 1322/2; 1

Query Match
Best Local Similarity 53.8%; Score 42; DB 2; Length 1777;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 NPDPYLYISKDV 15
| | | | | | | | | |
DB 1284 NPDSYGLNDVNG 1296

Search completed: June 26, 2003, 00:52:21
Job time : 3.47059 secs

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:19 ; Search time 1.58824 Seconds
(without alignment)

391.721 million cell updates/sec

Title: US-09-980-370-1

Sequence: 1 ALNPDNYLISKDVVG 15

```
Scoring table:  BLOSUM62
                  Canoe 10 0
                  Canoe 0 5
```

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1128922

Minimum DB seq length: 0

Maximum no. of seq. analyzed: 200000000

Maximum Match 1008

Database : Swjapbprot AN.v

pred NO is the number of results predicted by chance to have a

score greater than or equal to the score of the result being
and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	889	1	PRPH_F08G1
2	78	100.0	889	1	HBAAL_C08G1
3	78	100.0	2528	1	PRPH_F08G1
4	45	57.7	358	1	CADH_M08A
5	45	57.7	358	1	CADH_M08A
6	41	53.6	313	1	GAS2_HUMAN
7	41	53.6	314	1	GAS2_MOUSE
8	41	53.6	462	1	TBC2_EUPRC
9	41	53.6	618	1	CHAO_CLOBI
10	40	51.3	186	1	LMH1_HUMAN
11	40	51.3	186	1	LMH1_MOUSE
12	39	50.0	274	1	NUCA_NANSP
13	39	50.0	368	1	ROX1_YEAST
14	39	50.0	471	1	CD36_HUMAN
15	38.5	49.4	302	1	LGUL_SCHPO
16	38.5	49.4	851	1	MYCL_F08P1
17	38.5	49.4	917	1	SC1_CAMLE
18	38	48.7	278	1	PRR8_MYELE
19	38	48.7	369	1	MYR1_HA3IN
20	38	48.7	418	1	LEU2_THRITN
21	38	48.7	459	1	MB33_ARATH
22	38	48.7	484	1	TRPQ_YEAST
23	38	48.7	613	1	KPEP_PHYBL
24	38	48.7	726	1	SNEI_CANAL
25	38	48.7	726	1	SNEI_CANAL
26	38	48.7	726	1	ANSA_C9RMA
27	38	48.7	726	1	ANSA_C9RMA
28	37.5	48.1	705	1	CBR1_BACSI
29	37	47.4	147	1	LECE_ANTCR
30	37	47.4	321	1	NADA_SUTRO
31	37	47.4	357	1	CADH_POEDE
32	37	47.4	429	1	DPD2_ORLAF
33	37	47.4	452	1	PR11_PLAFLK
34	37	47.4	452	1	PR11_PLAFLK
35	37	47.4	452	1	PR11_PLAFLK
36	37	47.4	452	1	PR11_PLAFLK
37	37	47.4	452	1	PR11_PLAFLK
38	37	47.4	452	1	PR11_PLAFLK
39	37	47.4	452	1	PR11_PLAFLK
40	37	47.4	452	1	PR11_PLAFLK
41	37	47.4	452	1	PR11_PLAFLK
42	37	47.4	452	1	PR11_PLAFLK
43	37	47.4	452	1	PR11_PLAFLK
44	37	47.4	452	1	PR11_PLAFLK
45	37	47.4	452	1	PR11_PLAFLK
46	37	47.4	452	1	PR11_PLAFLK
47	37	47.4	452	1	PR11_PLAFLK
48	37	47.4	452	1	PR11_PLAFLK
49	37	47.4	452	1	PR11_PLAFLK
50	37	47.4	452	1	PR11_PLAFLK
51	37	47.4	452	1	PR11_PLAFLK
52	37	47.4	452	1	PR11_PLAFLK
53	37	47.4	452	1	PR11_PLAFLK
54	37	47.4	452	1	PR11_PLAFLK
55	37	47.4	452	1	PR11_PLAFLK
56	37	47.4	452	1	PR11_PLAFLK
57	37	47.4	452	1	PR11_PLAFLK
58	37	47.4	452	1	PR11_PLAFLK
59	37	47.4	452	1	PR11_PLAFLK
60	37	47.4	452	1	PR11_PLAFLK
61	37	47.4	452	1	PR11_PLAFLK
62	37	47.4	452	1	PR11_PLAFLK
63	37	47.4	452	1	PR11_PLAFLK
64	37	47.4	452	1	PR11_PLAFLK
65	37	47.4	452	1	PR11_PLAFLK
66	37	47.4	452	1	PR11_PLAFLK
67	37	47.4	452	1	PR11_PLAFLK
68	37	47.4	452	1	PR11_PLAFLK
69	37	47.4	452	1	PR11_PLAFLK
70	37	47.4	452	1	PR11_PLAFLK
71	37	47.4	452	1	PR11_PLAFLK
72	37	47.4	452	1	PR11_PLAFLK
73	37	47.4	452	1	PR11_PLAFLK
74	37	47.4	452	1	PR11_PLAFLK
75	37	47.4	452	1	PR11_PLAFLK
76	37	47.4	452	1	PR11_PLAFLK
77	37	47.4	452	1	PR11_PLAFLK
78	37	47.4	452	1	PR11_PLAFLK
79	37	47.4	452	1	PR11_PLAFLK
80	37	47.4	452	1	PR11_PLAFLK
81	37	47.4	452	1	PR11_PLAFLK
82	37	47.4	452	1	PR11_PLAFLK
83	37	47.4	452	1	PR11_PLAFLK
84	37	47.4	452	1	PR11_PLAFLK
85	37	47.4	452	1	PR11_PLAFLK
86	37	47.4	452	1	PR11_PLAFLK
87	37	47.4	452	1	PR11_PLAFLK
88	37	47.4	452	1	PR11_PLAFLK
89	37	47.4	452	1	PR11_PLAFLK
90	37	47.4	452	1	PR11_PLAFLK
91	37	47.4	452	1	PR11_PLAFLK
92	37	47.4	452	1	PR11_PLAFLK
93	37	47.4	452	1	PR11_PLAFLK
94	37	47.4	452	1	PR11_PLAFLK
95	37	47.4	452	1	PR11_PLAFLK
96	37	47.4	452	1	PR11_PLAFLK
97	37	47.4	452	1	PR11_PLAFLK
98	37	47.4	452	1	PR11_PLAFLK
99	37	47.4	452	1	PR11_PLAFLK
100	37	47.4	452	1	PR11_PLAFLK

54	37	47.4	49.2	VAB2-ACENT	Q38660	acetabular
53	37	47.4	50.9	HAI1-HUMAN	P38567	homo sapiens
52	37	47.4	67.7	NSP_KILDA	P23211	Kluyveromyces
51	37	47.4	68.1	YDBE-SCHRO	Q92359	Schizosaccharomyces
49	37	47.4	76.1	AFG3-YEAST	P39325	Saccharomyces
48	37	47.4	81.8	PSGA-SIN3	Q57105	Synthetic
47	37	47.4	82.1	SPK1-YEAST	P23216	Saccharomyces
46	37	47.4	83.1	NAP1-PARP	Q6350	paracoccus
45	37	47.4	84.6	NRT-DROME	P33554	Drosophila
44	36	46.2	179	VSG1-TRIT	P20849	tritus moria

ALIGNMENTS

RESULT 1

ID	PRTH_PORGI	STANDARD;	PRT;	989 AA.
10	PRTH_PORGI	STANDARD;	PRT;	989 AA.
11	PRTH_PORGI	STANDARD;	PRT;	989 AA.

DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel 32, Last sequence update)

16-OCT-2001 (Rel. 40, last annotation update;
Protease orth (EC 3.4.23.7)).

GN
PRTH.
GN
porphyromonas gingivalis (Bacteroides gingivalis)

OC Bacteroides; Bacteroidetes; Bacteroides; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.

```

Ox NCBI_18X10=63/;
RN [1]

```

KF SEQUENCE FROM N.A.
RC STRAIN-W83;

RA Fletcher H.M., Schenkein H.A., Macrina F.L.;

RT Porphyromonas gingivalis.";

RN [2]

RA Fletcher H.M., Schenkeln H.A., Macrina F.L.

-1- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENABLE ACTIVATION TO ENHANCE COMPLEMENT-MEDIATED KILLING THE

CC IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS

CC -1- SUBCELLULAR LOCATION: IN MEMBRANE VESICLES.
CC -1- SIMILARITY: BELONGS TO HERPESGAE FAMILY C25

 CC This Swiss-Born entry is convicted. It is produced
 CC

CC between the Swiss Institute of Bioinformatics
CC the European Bioinformatics Institute. There a

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC -----
DR -----
EMBL: F27483: AAA51298.1: -
```

```
DR      HSSP; P23882; 1FMT.  
DR      MEROPS; C25.001; -.
```

```
DR InterPro; IPR001705; peptidase_c23.  
DR InterPro; IPR002376; formyl_transf.
```

DR Pfam; PF00551; 101myl_c1rnsl; 1.
DR Pfam; PF01364; peptidase_C25; 1.

KM	HYDROLASE,	MILK	PROTEASE,	REPEAT,	VARIANCE.
FT	REPEAT	270	323		

```

SQ SEQUENCE 989 AA; 110238 MW; FA85FE8A3AC8944C CRC64;

```

```
Query Match      100.08; Score 78; DB 1; Length 989;
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY      1 ALPNPNYLISKDVTC 15
DB      149 ALPNPNYLISKDVTC 163

RESULT 2
HAGA_PORCI STANDARD: PRT: 2628 AA.
AC 051845-2000 (rel. 39, Created)
AD 051845-2000 (rel. 39, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DE Hemagglutinin A precursor.
GN HAGA.
OS Porphyromonas gingivalis (Bacteroides gingivalis),
OC Bacteria; Bacteroidetes; Bacteroidales; Bacteroidales;
OC Porphyromonadaceae; Porphyromonadaceae;
OX NCBI_TaxID=837;

RP SEQUENCE FROM N.A.
RX STRAIN=381;
RC MEDLINE=97047672; PubMed=8926061;
RT "The hemagglutinin (HAGA) of Porphyromonas gingivalis 381
RT contains four large, contiguous, direct repeats."
RL Infect. Immun. 64:4000-4007(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC THIS SWISS-PROT entry is a copy of the protein structure
CC between the Swiss Institute of Bioinformatics and the EMBL
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: U01807; A01128.1;
DR InterPro: IPR001769; Peptidase_C25;
KM Hemagglutinin; Virulence; Hydrolyase; Signal; Repeat.
FT SIGNAL 25 2628
FT DOMAIN 25 539 HEMAGGLUTININ A.
FT DOMAIN 540 995 PEPTIDASE C25-LIKE 1.
FT DOMAIN 996 1451 PEPTIDASE C25-LIKE 2.
FT DOMAIN 1452 1907 PEPTIDASE C25-LIKE 3.
FT DOMAIN 2074 2628 PEPTIDASE C25-LIKE 4.
SO SEQUENCE 2628 AA; 2683324 MW; 61C4DB32540C950A CRC64;

Query Match
Best Local Similarity 100.0%; Score 78; DB 1; Length 2628;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ALPNPNYLISKDVTC 15
DB      744 ALPNPNYLISKDVTC 758

RESULT 3
TREG_COEHE STANDARD: PRT: 265 AA.
AD 01-FEB-1995 (rel. 31, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DE Tubulin gamma chain (gamma tubulin) (Fragmant).
OS Cochliobolus heterostrophus (Drechslera maydis).
OC Cochliobolales; Pleosporomycetes; Pleosporomycetes;
OC Pleosporales; Pleosporomycetes; Cochliobolales.
OX NCBI_TaxID=5016;

RP SEQUENCE FROM N.A.
RA Parkinson C., Luo H., Knight A., Ahlquist J., Perlman M.H.;
RA Submitted (GNG-1993) to the EMBL/GenBank/DBD databases.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC THIS SWISS-PROT entry is a copy of the protein structure
CC between the Swiss Institute of Bioinformatics and the EMBL
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: X74455; CA52464.1;
DR InterPro: IPR000217; Tubulin;
DR InterPro: IPR003008; Tubulin_FtsZ.
KM PROSITE: P800227; TUBULIN; 1.
KM Microtubules; GTP-binding.
FT NP_BIND 77 83
FT NON_TER 265 265 GTP (POTENTIAL).
SO SEQUENCE 265 AA; 29567 MW; ASDAC23E7D62DC6 CRC64;

Query Match
Best Local Similarity 61.3%; Score 45; DB 1; Length 265;
Matches 0; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY      3 NPNPNYLISKDVTC 15
DB      22 NPNPNYLISKDVTC 34

RESULT 4
CADL_MEDSA STANDARD: PRT: 358 AA.
AC 01-JUL-1993 (rel. 26, Created)
DT 01-JUL-1993 (rel. 26, Last sequence update)
DE Cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) (CAD).
GN CAD2.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Fabaceae; Papilionoideae; Trifoliales; Medicago.
OX NCBI_TaxID=3879;

RP SEQUENCE FROM N.A.
RA van Doorslaere J., Baucher M., Feuillet C., Boudet A.M.;
RA "The cinnamyl alcohol dehydrogenase cDNAs from two important
RA economic species: alfalfa and poplar. Demonstration of a high homology
RA of the gene within angiosperms."
RL Plant Physiol. Biochem. 33:105-109(1995).
RN [2]
RN SEQUENCE FROM N.A. AND CHARACTERIZATION.
RN MEDLINE=20044095; PubMed=1057844;
RN WILL E.M., Abrahams S., Hayes C.M., Jenkins C.L., Watson J.M.;
RN "Molecular characterization and expression of a wound-inducible cDNA
RN encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne
RN (Medicago sativa L.)."
RN Plant Mol. Biol. 11:279-281(1990)
RN -1- PRELIMINARY SYNTHESIS SPECIFIC FOR PRODUCTION OF LIGNIN
RN MONOMERS. IT ACTS ON CONIFERYL-, SINAPYL-, 4-CUMARYL- AND
RN CINNAMYL-ALCOHOL.

```


CC -1- CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) -> cinnamaldehyde +
 CC NADPH.
 CC -1- COFACTOR: ZINC (BY SIMILARITY).
 CC -1- PATHWAY: Lipid synthesis.
 CC -1- TISSUE SPECIFICITY: MOST ACTIVELY EXPRESSED IN STEM, HYPOCOTYL AND
 CC ROOT TISSUE.
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z19573; CAA79625.1; -
 CC EMBL: AF083332; AAC35845.1; -
 CC PIR: S31572; S31572.
 CC InterPro: IPR002328; ADH_zinc.
 CC InterPro: IPR02085; Adh_zn_family.
 CC Pfam: PF00107; adh_zinc.1.
 CC PROSITE: PS00059; ADH_ZINC.1.
 CC Oxidoreductase; NADP; zinc; Lipid biosynthesis.
 CC METAL 48 48 ZINC (CATALYTIC) (BY SIMILARITY).
 CC METAL 70 70 ZINC (CATALYTIC) (BY SIMILARITY).
 CC METAL 101 101 ZINC (SECOND ATOM) (BY SIMILARITY).
 CC METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).
 CC METAL 107 107 ZINC (SECOND ATOM) (BY SIMILARITY).
 CC METAL 115 115 ZINC (SECOND ATOM) (BY SIMILARITY).
 CC METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).
 CC SEQUENCE 358 AA; 38948 MW; FBA609408D01B956 CRC64;
 CC
 CC Query Match 57.7% Score 45; DB 1; Length 358;
 CC Best Local Similarity 57.1% Pred. No. 2.9;
 CC Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 CC
 CC Oy 2 LNPNDYLSMDYTG 15
 CC Db 225 LGADNYLVSDTVG 238
 CC
 CC RESULT 5
 CC GAS2_HUMAN STANDARD; PRT; 313 AA.
 CC ID GAS2_HUMAN
 CC AC 043903;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Growth-arrest-specific protein 2 (GAS-2).
 CC GN GAS2.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Kidney;
 CC RX MEDLINE=98190528; PubMed=9521882;
 CC RA Collavin L., Buzzaï M., Saccone S., Bernard L., Federico C.,
 CC della Valle G., Brancolini C., Schneider C.;
 CC "cDNA characterization and chromosome mapping of the human GAS2
 CC gene".
 CC Genomics 48:265-269(1998).
 CC -1- FUNCTION: MAY PLAY A ROLE IN APOPTOSIS BY ACTING AS A CELL DEATH
 CC SUBSTRATE FOR CASPASES. IS CLEAVED DURING APOPTOSIS AND THE
 CC CLEAVED FORM INDUCES DRAMATIC REARRANGEMENTS OF THE ACTIN
 CC CYTOSKELETON AND POTENT CHANGES IN THE SHAPE OF THE AFFECTED
 CC CELLS. MAY BE INVOLVED IN THE MEMBRANE RUFLING PROCESS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: COMPONENT OF THE MICROFILAMENT SYSTEM.
 CC COLOCALIZES WITH ACTIN FIBERS AT THE CELL BORDER AND ALONG THE

CC STRESS FIBERS IN GROWTH-ARRESTED FIBROBLASTS. MAINLY
 CC MEMBRANE-ASSOCIATED. WHEN HYPERPHOSPHORYLATED, ACCUMULATES AT
 CC MEMBRANE RUFLINGS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: UNICITOUSLY EXPRESSED WITH HIGHEST LEVELS IN
 CC LIVER, LUNG, AND KIDNEY. NOT FOUND IN SPLEEN.
 CC -1- DEVELOPMENTAL STAGE: SPECIFICALLY EXPRESSED AT GROWTH ARREST.
 CC -1- PTM: CLEAVED, DURING APOPTOSIS, ON A SPECIFIC ASPARTIC RESIDUE BY
 CC CASPASES.
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES DURING THE G0-G1 TRANSITION
 CC PHASE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GAS2 FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U95032; AAC52058.1; -
 CC EMBL: HGNC:4167; GAS2.
 CC Gene: HGNC:4167; GAS2.
 CC MIM: 602835; -
 CC InterPro: IPR001715; Calponin-like.
 CC InterPro: IPR003108; GAS2.
 CC Pfam: PF02187; GAS2.1.
 CC SMART: SM00033; CH.1.
 CC SMART: SM00243; GAS2.1.
 CC PROSITE: PS00021; CH.1.
 CC KW Growth arrest; Phosphorylation; Apoptosis; Cell cycle;
 CC Cytoskeleton.
 CC FT DOMAIN 34 156 CH.
 CC FT SITE 278 279 CLEAVAGE (BY A CASPASE DURING
 CC APOPTOSIS).
 CC SEQUENCE 313 AA; 34945 MW; A5FBA611A0C8B36F CRC64;
 CC
 CC Query Match 52.6% Score 41; DB 1; Length 313;
 CC Best Local Similarity 75.0% Pred. No. 12;
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Oy 2 LNPNDYLI 9
 CC Db 294 NNPNDYLV 301
 CC
 CC RESULT 6
 CC GAS2_MOUSE STANDARD; PRT; 314 AA.
 CC ID GAS2_MOUSE
 CC AC P11862;
 CC DT 01-OCT-1989 (Rel. 12, Created)
 CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Growth-arrest-specific protein 2 (GAS-2).
 CC GN GAS2 OR GAS-2.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RC MEDLINE=88311081; PubMed=3409319;
 CC RA Schneider C., King R.M., Philipson L.;
 CC "Genes specifically expressed at growth arrest of mammalian cells".
 CC Cell 54:787-793(1988).
 CC [2]
 CC RP CHARACTERIZATION.
 CC RX MEDLINE=92299682; PubMed=1607387;
 CC RA Brancolini C., Bottega S., Schneider C.;
 CC "GAS2, a growth arrest-specific protein, is a component of the
 CC microfilament network system".
 CC J. Cell Biol. 117:1251-1261(1992).
 CC [3]

RP PHOSPHORYLATION, AND SUBCELLULAR LOCATION.
 RX MEDLINE-94165141; PubMed-812096;
 RA Brancolini C., Schneider C.;
 RT "phosphorylation of the growth arrest-specific protein Gas2 is
 RT coupled to actin rearrangements during G0->G1 transition in
 RT NIH 3T3 cells";
 RL J. Cell Biol. 124:743-756(1994).
 RN [4]
 RP PROCESSING
 RX MEDLINE-96080153; PubMed-7489707;
 RA Brancolini C., Benedetti M., Schneider C.;
 RT "Microfilament reorganization during apoptosis: The role of Gas2, a
 RT possible substrate for ICE-like proteases";
 RL EMBO J. 14:5179-5190(1995).
 RN [5]
 RP FUNCTION AND DEVELOPMENTAL STAGE
 RX MEDLINE-99160524; PubMed-10049561;
 RA Lee K.K., Tang M.K., Yew D.T., Chow P.H., Yee S.P., Schneider C.,
 RA Brancolini C.;
 RT "gas2 is a multifunctional gene involved in the regulation of
 RT apoptosis and chondrogenesis in the developing mouse limb.";
 RL dev. Biol. 207:14-25(1999).
 CC -1- FUNCTION: MAY PLAY A ROLE IN APOPTOSIS BY ACTING AS A CELL DEATH
 CC SUBSTRATE FOR CASPASES. IS CLEAVED DURING APOPTOSIS AND THE
 CC CLEAVED FORM INDUCES DRAMATIC REARRANGEMENTS OF THE ACTIN
 CC CYTOSKELETON AND POTENT CHANGES IN THE SHAPE OF THE AFFECTED
 CC CELLS. MAY PLAY A ROLE IN CHONDROCYTE PROLIFERATION AND
 CC DIFFERENTIATION, AND IN LIMB MYOGENESIS. MAY BE INVOLVED IN THE
 CC REGULATION OF THE APOPTOSIS IN THE INTERDIGITAL TISSUES OF THE
 CC DEVELOPING HINDLIMB. MAY BE INVOLVED IN THE MEMBRANE RUFLING
 CC PROCESS.
 CC -1- SUBCELLULAR LOCATION: COMPONENT OF THE MICROFILAMENT SYSTEM.
 CC COLOCALIZES WITH ACTIN FIBERS AT THE CELL BORDER AND ALONG THE
 CC STRESS FIBERS IN GROWTH-ARRESTED FIBROBLASTS. MAINLY
 CC MEMBRANE-ASSOCIATED. WHEN HYPERPHOSPHORYLATED, ACCUMULATES AT
 CC MEMBRANE RUFLINGS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. HIGHEST LEVELS IN
 CC LIVER, LUNG AND KIDNEY. IN THE EMBRYO STRONGLY EXPRESSED IN
 CC REGIONS THAT UNDERGO EXTENSIVE APOPTOSIS, SUCH AS THE
 CC INTERDIGITAL TISSUES, THE CRANIOFACIAL MESenchyme AND THE
 CC CARTRIDGE OF THE LIMBS.
 CC -1- DEVELOPMENTAL STAGE: AT EMBRYONIC DAYS E11.5 AND E13.5 STRONGLY
 CC EXPRESSED IN THE SOFT CONNECTIVE TISSUE OF THE FACE AND TRUNK, AND
 CC IN THE INVERTEBRAL TISSUES. LOW LEVELS ARE FOUND IN BRAIN AND
 CC NEURAL TUBE. LOW LEVELS ARE FOUND IN DAY E13.5 LUNG, KIDNEY, EYE
 CC LENS AND IN VERTEBRAL CARTILAGE LOCATED CRANIALLY. IN DAY E11.5
 CC HINDLIMBS WEAKLY EXPRESSED BY THE MESENCHYMAL CELLS SURROUNDING
 CC THE PERSPECTIVE CARTILAGE-FORMING REGIONS. IN DAY E12.5 HINDLIMBS
 CC STRONGLY EXPRESSED BY CELLS ENVELOPING THE CHONDROGENIC PRIMORDIA
 CC OF THE DIGITS, METATARSALS, TIBIA, AND FEMUR, AND THE SOFT
 CC CONNECTIVE TISSUE IN THE INTERDIGITAL TISSUES. IN DAY E13.5
 CC HINDLIMBS EXPRESSION IS MAINTAINED IN THE INTERDIGITAL TISSUES
 CC LOCATED PROXIMALLY AND IS FOUND IN SOME CHONDROCYTES IN THE
 CC STYLOPOD AND IN MESENCHYMAL CELLS SURROUNDING THE CARTILAGE IN THE
 CC AUTOPOD AND ZYGOPOD. IN DAY 13.5 FORELIMB STRONGLY EXPRESSED IN
 CC THE PRE-HYPERTROPHIC AND HYPERTROPHIC REGIONS OF THE HUMERUS.
 CC RADIUS, AND ULNA. EXPRESSION IN HYPERTROPHIC CHONDROCYTES IS
 CC MAINTAINED AT DAY E14.5 AND IS NOT DETECTABLE AT DAY E15.5. AT DAY
 CC 14.5 ALSO EXPRESSED BY CHONDROCYTES IN THE CARTILAGE FORMING THE
 CC CARPALS AND TARSALS AND BY MESENCHYMAL CELLS IN THE PROCESS OF
 CC CONDENSING TO FORM TENDONS. IN DAY E13.5 HINDLIMBS EXPRESSED IN
 CC SOME MYOBLASTS IN THE PROXIMAL MYOGENIC REGION. IN OLDER LIMBS
 CC EXPRESSION IS MAINTAINED IN THE MYOTUBULES.
 CC -1- INDUCTION: DOWN-REGULATED BY MITOGENS.
 CC -1- PTM: CLEAVED DURING APOPTOSIS, ON A SPECIFIC ASPARTIC RESIDUE BY
 CC CASPASES (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES DURING THE G0-G1 TRANSITION
 CC PHASE.
 CC -1- SIMILARITY: BELONGS TO THE GAS2 FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-2 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M21828; AAA37660.1; -
 CC DR PIR: A31580; A31590
 CC DR MGD: MGI:95657; Gas2
 CC DR InterPro: IPR001715; Calponin-like.
 CC DR InterPro: IPR003108; GAS2.
 CC DR Pfam: PF02187; GAS2; 1.
 CC DR SMART: SM00033; CH; 1.
 CC DR SMART: SM00243; GAS2; 1.
 CC DR PROSITE: PS50021; CH; 1.
 CC KW Growth arrest; Phosphorylation; Apoptosis; Cell cycle;
 CC KM Cytoskeleton.
 CC FT DOMAIN 35 157 CH.
 CC FT SITE 279 280 CLEAVAGE (BY A CASPASE DURING
 CC FT MUTAGEN 279 279 APOPTOSIS).
 CC FT MUTAGEN D->A: ABOLISHES PROTEOLYTIC PROCESSING.
 CC SQ SEQUENCE 314 AA; 34900 MW; 7F2CC704B4057FAC CRC64;
 CC
 CC Query Match 52.68; Score 41; DB 1; Length 314;
 CC Best Local Similarity 75.08; Pred. No. 12;
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 LNPNYIL 9
 CC DB 295 MNPNYLV 302
 CC
 CC RESULT 7
 CC TBG2_EUPCR STANDARD: PRT: 462 AA.
 CC AC P54404;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Tubulin gamma-2 chain (Gamma-2 tubulin).
 CC OS Euplates crassus.
 CC OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
 CC OC Euplotida; Euplotidae; Moneuplotes.
 CC OX NCBI_TaxID=5936;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE-98192540; PubMed-9524221;
 CC RA Tan M., Heckmann K.;
 CC RT "The two gamma-tubulin-encoding genes of the ciliate Euplates crassus
 CC differ in their sequences, codon usage, transcription initiation
 CC sites and poly(A) addition sites";
 CC RL Gene 210:53-60(1998).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
 CC GAMMA TUBULIN IS FOUND AT MICROTUBULE ORGANIZING CENTERS (MTOC)
 CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
 CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X85235; CA549490.1; -
 CC DR EMBL: Y09551; CA70742.1; -
 CC DR InterPro: IPR000217; Tubulin.
 CC DR InterPro: IPR003008; Tubulin_FtsZ.
 CC DR Pfam: PF00091; tubulin_1.
 CC DR PRINTS: PRO1161; TUBULIN.

DR PROSITE: PS00227; TUBULIN; 1.
 KW Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 142 148 CTP (POTENTIAL)
 SO SEQUENCE 462 AA: 52117 MW: 56387559AC4846 CRC64;
 Query Match 52.6%; Score 41; DB 1; Length 462;
 Best Local Similarity 46.2%; Pred. No. 19;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 3 NPDNYLISKDVTG 15
 11:1:11:1:1
 Db 87 NPENFVSKTMDG 99

RESULT 8
 CHAA_CLOBI STANDARD: PRT: 618 AA.
 ID CHAA_CLOBI
 AC 005102;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pesticidal crystal-like protein cry17Aa (insecticidal toxin
 DE CRYXVIIAa) (Cpm72 mosquitocidal toxin).
 GN CRY17AA OR CRYXVIIAa) OR CBM72.
 OS Clostridium bifermentans.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 CC Clostridium.
 CC NCBI_Taxid=1490;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH18 / subsp. Malaysia;
 RX MEDLINE=98267211; PubMed=9602158;
 RA Barloy F., Lecadet M.-M., Delecluse A.;
 RT Cloning and sequencing of three new putative toxin genes from
 RL Clostridium bifermentans CH18.";
 Gene 211:293-299(1998)
 CC -1- FUNCTION: NOT SIGNIFICANTLY TOXIC TO MOSQUITO LARVAE MAY INCREASE
 THE TOXICITY OF THE PESTICIDIAL CRYSTAL-LIKE PROTEIN CRY16AA
 (CBM71).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- DEVELOPMENTAL STAGE: PRODUCED DURING SPOULATION.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X99478; CAA67841.1;
 DR InterPro: IPR001178; Endotoxin.
 DR Pfam: PF00555; endotoxin; 1.
 KW Toxin; Sporulation.
 FT DOMAIN 174 179 POLY-LEU.
 FT SEQUENCE 618 AA: 71651 MW: 1B9C5251B9AD329D CRC64;
 SO SEQUENCE 618 AA: 71651 MW: 1B9C5251B9AD329D CRC64;
 Query Match 52.6%; Score 41; DB 1; Length 618;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 ALNPDNYLISKD 12
 11:1:11:1:1
 Db 436 SISPENYLFDKD 447

RESULT 9
 LMB1_HUMAN STANDARD: PRT: 1786 AA.
 ID LMB1_HUMAN
 AC P07942;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin beta-1 chain precursor (Laminin B1 chain).
 GN LMB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368768; PubMed=1975589;
 RA Voelckenhof R., Chow L.T., Tryggvason K.;
 RT Structure of the human laminin B1 chain gene.";
 RL J. Biol. Chem. 265:15611-15616(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87280097; PubMed=3611077;
 RA Pikkarienen T., Eddy R., Fukushima Y., Byers M., Shows T.,
 RA Pihlajaniemi T., Saraste M., Tryggvason K.;
 RT Human laminin B1 chain. A multidomain protein with gene (LMB1)
 RL locus in the q22 region of chromosome 7.";
 J. Biol. Chem. 262:10454-10462(1987).
 RN [3]
 RP SEQUENCE OF 1276-1709 FROM N.A.
 RX MEDLINE=88021029; PubMed=3661559;
 RA Jaye M., Modi W.S., Ricca G.A., Mudd R., Chiu I.M., O'Brien S.J.,
 RA Drohan W.N.;
 RT "Isolation of a cDNA clone for the human laminin-B1 chain and its
 RL gene localization.";
 RL Am. J. Hum. Genet. 41:605-615(1987).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
 CC 2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 CC COMPONENT).
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M61951; AAA59486.1;
 DR EMBL: M58147; AAA59486.1; JOINED.
 DR EMBL: M61917; AAA59486.1; JOINED.
 DR EMBL: M61918; AAA59486.1; JOINED.
 DR EMBL: M61921; AAA59486.1; JOINED.
 DR EMBL: M61922; AAA59486.1; JOINED.
 DR EMBL: M61923; AAA59486.1; JOINED.
 DR EMBL: M61924; AAA59486.1; JOINED.
 DR EMBL: M61925; AAA59486.1; JOINED.
 DR EMBL: M61926; AAA59486.1; JOINED.
 DR EMBL: M61927; AAA59486.1; JOINED.
 DR EMBL: M61928; AAA59486.1; JOINED.
 DR EMBL: M61929; AAA59486.1; JOINED.
 DR EMBL: M61930; AAA59486.1; JOINED.
 DR EMBL: M61931; AAA59486.1; JOINED.

DR EMBL: M61932: AAA59486.1: JOINED.
 DR EMBL: M61933: AAA59486.1: JOINED.
 DR EMBL: M61934: AAA59486.1: JOINED.
 DR EMBL: M61935: AAA59486.1: JOINED.
 DR EMBL: M61936: AAA59486.1: JOINED.
 DR EMBL: M61938: AAA59486.1: JOINED.
 DR EMBL: M61939: AAA59486.1: JOINED.
 DR EMBL: M61940: AAA59486.1: JOINED.
 DR EMBL: M61941: AAA59486.1: JOINED.
 DR EMBL: M61942: AAA59486.1: JOINED.
 DR EMBL: M61943: AAA59486.1: JOINED.
 DR EMBL: M61944: AAA59486.1: JOINED.
 DR EMBL: M61945: AAA59486.1: JOINED.
 DR EMBL: M61947: AAA59486.1: JOINED.
 DR EMBL: M61948: AAA59486.1: JOINED.
 DR EMBL: M61949: AAA59486.1: JOINED.
 DR EMBL: M61950: AAA59486.1: JOINED.
 DR EMBL: M55370: AAA59485.1: JOINED.
 DR EMBL: M55378: AAA59485.1: JOINED.
 DR EMBL: M55365: AAA59485.1: JOINED.
 DR EMBL: M55371: AAA59485.1: JOINED.
 DR EMBL: M55372: AAA59485.1: JOINED.
 DR EMBL: M55373: AAA59485.1: JOINED.
 DR EMBL: M55374: AAA59485.1: JOINED.
 DR EMBL: M55375: AAA59485.1: JOINED.
 DR EMBL: M55376: AAA59485.1: JOINED.
 DR EMBL: M55344: AAA59485.1: JOINED.
 DR EMBL: M55345: AAA59485.1: JOINED.
 DR EMBL: M55347: AAA59485.1: JOINED.
 DR EMBL: M55348: AAA59485.1: JOINED.
 DR EMBL: M55349: AAA59485.1: JOINED.
 DR EMBL: M55351: AAA59485.1: JOINED.
 DR EMBL: M55352: AAA59485.1: JOINED.
 DR EMBL: M55353: AAA59485.1: JOINED.
 DR EMBL: M55355: AAA59485.1: JOINED.
 DR EMBL: M55356: AAA59485.1: JOINED.
 DR EMBL: M55357: AAA59485.1: JOINED.
 DR EMBL: M55358: AAA59485.1: JOINED.
 DR EMBL: M55359: AAA59485.1: JOINED.
 DR EMBL: M55360: AAA59485.1: JOINED.
 DR EMBL: M55361: AAA59485.1: JOINED.
 DR EMBL: M55362: AAA59485.1: JOINED.
 DR EMBL: M55363: AAA59485.1: JOINED.
 DR EMBL: M55364: AAA59485.1: JOINED.
 DR EMBL: M55366: AAA59485.1: JOINED.
 DR EMBL: M55367: AAA59485.1: JOINED.
 DR EMBL: M55368: AAA59485.1: JOINED.
 DR EMBL: M55369: AAA59485.1: JOINED.
 DR EMBL: M61916: AAA59482.1: -.
 DR EMBL: M20206: AAA59487.1: -.
 DR PIR: S13547: MMHUB1.
 DR HSP: P02468: 1KLO.
 DR Genew: HGNC:6486: LAMB1.
 DR MIM: 150240: -.
 DR InterPro: IPR000561: EGF-like.
 DR InterPro: IPR001886: Lamnt.
 DR InterPro: IPR002049: Laminin_EGF.
 DR Pfam: PF00053: Laminin_EGF_13.
 DR Pfam: PF00055: Laminin_Nterm_1.
 DR PRINTS: PR00011: EGF_LAMININ.
 DR ProDom: PD002082: Lamnt_1.
 DR SMART: SM00180: EGF_Lam_11.
 DR SMART: SM00136: Lamnt_1.
 DR PROSITE: PS00022: EGF_1_9.
 DR PROSITE: PS01186: EGF_2_2.
 DR PROSITE: PS01248: LAMININ_Type_EGF_11.
 DR GlycoProtDB: Basement membrane: Extracellular matrix: Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNL 1 21
 FT CHAIN 22 1786 LAMININ BETA-1 CHAIN.

FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
 FT DOMAIN 335 397 LAMININ EGF-LIKE 2.
 FT DOMAIN 398 437 LAMININ EGF-LIKE 3.
 FT DOMAIN 438 509 LAMININ EGF-LIKE 4.
 FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 541 771 LAMININ DOMAIN IV.
 FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
 FT DOMAIN 1179 1397 DOMAIN II.
 FT DOMAIN 1398 1430 DOMAIN I.
 FT DOMAIN 1431 1786 DOMAIN I ALPHA.
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1353 1388 COILED COIL (POTENTIAL).
 FT DOMAIN 1442 1781 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 448 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 869 883 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.

Query Match 51.3%. Score 40: DB 1; Length 1786;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPNYLISKDVT 14
 1111:1111
 DB 89 LNPDSHLIENVY 101

RESULT 10
 ID LMB1_MOUSE STANDARD: PRT: 1786 AA.
 AC P02469:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin beta-1 chain precursor (Laminin B1 chain).
 GN LAMB1-1 OR LAMB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090.

RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87147212; PubMed=3493487;
 RT Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;
 "Sequence of the cDNA encoding the laminin B1 chain reveals a
 RT multidomain protein containing cysteine-rich repeats."
 RL Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).
 RP [2]
 RA SEQUENCE OF 1292-1786 FROM N.A.
 RX MEDLINE=85051302; PubMed=6209134;
 RT Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
 "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
 RT coiled-coil alpha-helix."
 RL EMBO J. 3:2355-2362(1984).
 RN [3]
 RP SEQUENCE OF 165-172; 539-547 AND 712-719.
 RC STRAIN-BALB/C; TISSUE=Endothelial cells;
 RX MEDLINE=97363207; PubMed=9219532;
 RT Frieser M., Noeckel H., Pausch F., Roeder C., Hahn A., Deutzmann R.,
 RA Sorokin L.M.;
 "Cloning of the mouse laminin alpha 4 cDNA. Expression in a subset of
 RT endothelium."
 RL Eur. J. Biochem. 246:727-735(1997).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
 CC 2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 CC COMPONENT).
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M15525; AAA39407.1; ALT_INIT.
 DR EMBL: X05212; CAA28839.1; -
 DR PIR: A26413; MMSB1..
 DR HSSP: P02468; IKLO.
 DR MGD: MGI:96743; LAMB1-1.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001886; LAMNT.
 DR pfam: PF00053; Laminin_EGF.
 DR pfam: PF00055; Laminin_EGF_13.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR PRODOM: PD002082; LAMNT; 1.
 DR SMART: SM00180; EGF_Lam; 11.
 DR SMART: SM00136; LAMNT; 1.
 DR PROSITE: PS00023; EGF_1; 9.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01248; LAMININ_Type_EGF; 11.
 DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 DR Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 KW SIGNAL
 FT CHAIN 1 21
 FT CHAIN 22 1786 LAMININ BETA-1 CHAIN.
 FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
 FT DOMAIN 335 397 LAMININ EGF-LIKE 2.

FT DOMAIN 398 457 LAMININ EGF-LIKE 3.
 FT DOMAIN 458 509 LAMININ EGF-LIKE 4.
 FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 541 772 LAMININ DOMAIN IV.
 FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
 FT DOMAIN 1179 1397 DOMAIN 11.
 FT DOMAIN 1398 1430 DOMAIN ALPHA.
 FT DOMAIN 1431 1786 DOMAIN I.
 FT DOMAIN 1787 1815 COILED COIL (POTENTIAL).
 FT DOMAIN 1816 1848 COILED COIL (POTENTIAL).
 FT DOMAIN 1849 1878 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 507 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 792 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.
 FT DISULFID 919 944 BY SIMILARITY.
 FT DISULFID 946 955 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 976 990 BY SIMILARITY.
 FT DISULFID 978 997 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1025 BY SIMILARITY.
 FT DISULFID 1084 1096 BY SIMILARITY.
 FT DISULFID 1086 1103 BY SIMILARITY.
 FT DISULFID 1105 1114 BY SIMILARITY.
 FT DISULFID 1117 1129 BY SIMILARITY.
 FT DISULFID 1132 1144 BY SIMILARITY.
 FT DISULFID 1134 1151 BY SIMILARITY.
 FT DISULFID 1153 1162 BY SIMILARITY.
 FT DISULFID 1165 1176 BY SIMILARITY.
 FT DISULFID 1179 1179 BY SIMILARITY.
 FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
 FT DISULFID 1185 1185 INTERCHAIN (PROBABLE).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC).
 FT CARBOHYD 519 519 N-LINKED (GLCNAC).
 FT CARBOHYD 677 677 N-LINKED (GLCNAC).
 FT CARBOHYD 1041 1041 N-LINKED (GLCNAC).
 FT CARBOHYD 1195 1195 N-LINKED (GLCNAC).
 FT CARBOHYD 1279 1279 N-LINKED (GLCNAC).
 FT CARBOHYD 1336 1336 N-LINKED (GLCNAC).

```

FT CARBOHYD 1343 1343 N-LINKED (GLICNAC. . .) (POTENTIAL)
FT CARBOHYD 1487 1487 N-LINKED (GLICNAC. . .) (POTENTIAL)
FT CARBOHYD 1533 1533 N-LINKED (GLICNAC. . .) (POTENTIAL)
FT CARBOHYD 1542 1542 N-LINKED (GLICNAC. . .) (POTENTIAL)
FT CARBOHYD 1643 1643 N-LINKED (GLICNAC. . .) (POTENTIAL)
FT CONFLICT 1531 1534 SGNA -> MEMP (IN REF. 2).
FT CONFLICT 1749 1749 D -> N (IN REF. 2).
SO SEQUENCE 1786 AA; 196904 MW; 846671B7BF1A474 CRC64;

Query Match 51.3%; Score 40; DB 1; Length 1786;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LNPDNYLISKDVT 14
Db 89 LNPDHSLIENVY 101

RESULT 11
PGKT_THEME STANDARD; PRT; 654 AA.
AC P36204; Q60031;
DT 01-JUN-1994 (Rel. 29, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bifunctional PCK/TIM (Includes: Phosphoglycerate kinase (EC 2.7.2.3);
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM)).
GN PCK/TPI OR TM0689.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=95163577; PubMed=7859734;
RA Schurig H., Beaucamp N., Ostendorf R., Jaenicke R., Adler E.,
RA Knowles J.R.;
RT "Phosphoglycerate kinase and triosephosphate isomerase from the
RT hyperthermophilic bacterium Thermotoga maritima form a covalent
RT bifunctional enzyme complex.";
RL EMBO J. 14:442-451(1995).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-326(1999).

RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF PKA AND REVISIONS.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=98046096; PubMed=9384563;
RA Auerbach G., Huber R., Graetlinger M., Zaiss K., Schurig H.,
RA Jaenicke R., Jacob U.;
RT "Closed structure of phosphoglycerate kinase from Thermotoga maritima
RT reveals the catalytic mechanism and determinants of thermal
RT stability.";
RL Structure 5:1475-1483(1997).

RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.85 ANGSTROMS) OF TIM.
RX MEDLINE=20058648; PubMed=10591103;
RA Mees D., Zeelen J.P., Thanki N., Beaucamp N., Alvarez M., Thi M.H.,
RA Bachmann J., Martal J.A., Wynn L., Jaenicke R., Wierenga R.K.;
RT "The crystal structure of triosephosphate isomerase (TIM) from
RT Thermotoga maritima: a comparative thermostability structural
RT analysis of ten different TIM structures.";

```

```

RL Proteins 37:441-453(1999).
CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycero-
CC phosphate.
CC -1- PATHWAY: Second phase of glycolysis; second step.
CC -1- SUBUNIT: MONOMER (PCK) AND HOMOTETRAMER (PCK-TIM).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC PHOSPHOGLYCERATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRIOSEPHOSPHATE ISOMERASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X75437; CAA53187.1; -
DR EMBL: L27492; AAB67520.1; -
DR EMBL: AE001741; AAD35771.1; -
DR PIR: S38439; S38439.
DR PDB: 1B9B; 01-JUN-00.
DR PDB: 1VEB; 17-JUN-98.
DR TIGR: TM0689; -
DR InterPro: IPR001576; PCK.
DR InterPro: IPR000652; Triophos_ismrse.
DR Pfam: PF001021; TIM; 1.
DR Pfam: PF00162; PCK; 1.
DR PRINTS: PR00477; PHGLYCKINASE.
DR ProDom: PD001005; Triophos_ismrse; 1.
DR TIGRFAMs: TIGR00419; tim; 1.
DR PROSITE: PS00111; POLYMERASE_KINASE; 1.
DR PROSITE: PS00171; TIM; 1.
KW Transferase; Kinase; Isomerase; Glycolysis; Gluconeogenesis;
KW Fatty acid biosynthesis; Pentose shunt; Multifunctional enzyme;
KW 3D-structure; Complete proteome.
FT DOMAIN 1 399 PHOSPHOGLYCERATE KINASE.
FT DOMAIN 2 400 654 TRIOSEPHOSPHATE ISOMERASE.
FT ACT_SITE 495 495 BY SIMILARITY.
FT ACT_SITE 567 567 BY SIMILARITY.
FT CONFLICT 213 213 D -> N (IN REF. 1).
FT CONFLICT 394 396 IAD -> MRI (IN REF. 1).
FT CONFLICT 626 626 K -> R (IN REF. 1).
FT CONFLICT 640 640 E -> Q (IN REF. 1).
SO SEQUENCE 654 AA; 71585 MW; 42358A4EF0C5E481 CRC64;

Query Match 50.6%; Score 39.5; DB 1; Length 654;
Best Local Similarity 38.9%; Pred. No. 49;
Matches 7; Conservative 5; Mismatches 3; Indels 3; Gaps 1;

OY 1 ALNPDNYLISKDVTG 15
Db 613 SIKPDNGLITVQKIDG 630

RESULT 12
NUCA_ANASP STANDARD; PRT; 274 AA.
AC P38446;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nuclease precursor (EC 3.1.30.-) (Endonuclease).
GN NUCA OR ALL7362.
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120alpha.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]

```


RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE-93125124; PubMed-1343821;
 RA Muro-Pastor A.M., Flores E., Herrero A., Molk C.P.;
 RT "Identification, genetic analysis and characterization of a
 RT sugar-non-specific nuclease from the cyanobacterium *Anabaena* sp. PCC
 RT 7120";
 RL Mol. Microbiol. 6:3021-3030(1992).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21595285; PubMed-11759840;
 RA Kameko T., Nakamura Y., Molk C.P., Kunitz T., Sasamoto S.,
 RA Metanaka A., Iritaguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium *Anabaena* sp. strain PCC 7120";
 RL DNA Res. 8:205-213(2001).
 CC -1- FUNCTION: CATALYZES THE DEGRADATION OF BOTH RNA AND DNA, HAS THE
 CC POTENTIAL TO ACT AS AN ENDONUCLEASE.
 CC -1- COFACTOR: REQUIRES A DIVALENT CATION FOR ACTIVITY; THE
 CC EFFECTIVENESS OF THE CATIONS ARE $Mg^{2+} > Mg^{2+} > Ca^{2+} = Co^{2+}$.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC OR LOOSELY ATTACHED TO THE
 CC CYTOSOL PLASMIC OR THE OUTER MEMBRANE.
 CC -1- PPM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO THE DNA/RNA NON-SPECIFIC ENDONUCLEASES
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64705; CAA45962.1; -;
 DR EMBL: AP003601; BAB77120.1; -;
 DR PIR: S28039; S28039;
 DR InterPro: IPR001604; Endonuclease.
 DR Pfam: PF01223; Endonuclease; 1.
 DR SMART: SM00477; NUC; 1.
 DR PROSITE: PS01070; NUCLEASE_NON_SPEC; 1.
 DR HydroLase; Nuclease; Endonuclease; Manganese; Signal;
 KW Periplasmic; Plasmid; Complete proteome.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 274 NUCLEASE.
 FT ACT_SITE 124 124 BY SIMILARITY.
 FT SEQUENCE 274 AA; 23669 MW; 7F14170915AE59BF CRC64;
 SO
 Query Match 50.0%; Score 39; DB 1; Length 274;
 Best Local Similarity 63.6%; Pred. No. 23;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 LNPNTYLSKDK 12
 Db 53 LTPDNTLVKVN 63
 RESULT 13
 ROX1_YEAST
 ID ROX1_YEAST STANDARD: PRT; 368 AA.
 AC P25042;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ROX1 repressor (Hypoxic function repressor) (Heme-dependent repression
 DE factor).
 GN ROX1 OR YPR065W OR YPR499.20.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 CC NCBI_TaxID=4932;

RM [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94019282; PubMed-8413209;
 RA Balasubramanian B., Lowry C.V., Zitomer R.S.;
 RT "The RoX1 repressor of the *Saccharomyces cerevisiae* hypoxic genes is
 RT a specific DNA-binding protein with a high-mobility-group motif";
 RL Mol. Cell. Biol. 13:6071-6078(1993).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-8288C / AB972;
 RP MEDLINE-97333271; PubMed-9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albertmann K., Allen E., Anstorge W.,
 RA Arujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Bostein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kaiman S., Klein R.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mirzaii S., Moestl D.,
 RA Mueller-Ruer S., Nemach A., Newkirk U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schaefer M., Scharte M.,
 RA Scherens B., Schramm S., Schroeder M., Sidcu A.M., Tietelin H.,
 RA Uristarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambut R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI";
 RL Nature 387:103-105(1997).
 RM [3]
 RP CHARACTERIZATION.
 RX MEDLINE-96174644; PubMed-8600445;
 RA di Flumeri C., Liston P., Acheson N.H., Keng T.;
 RT "The HMG domain of the RoX1 protein mediates repression of HEM13
 RT through overlapping DNA binding and oligomerization functions";
 RL Nucleic Acids Res. 24:808-815(1996).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT REPRESSSES THE EXPRESSION OF
 CC HEM13, COX5B, ANB1, CYC7 OR ANC3 (HYPOXIC FUNCTION). BINDS TO THE
 CC DNA SEQUENCE 5'-RRRTAACAGAG-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- INDUCTION: BY HEME.
 CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X60458; CAA42991.1; -;
 DR EMBL: 249219; CAA89183.1; -;
 DR EMBL: 271255; CAA94973.1; -;
 DR PIR: S17015; S17015;
 DR HSP: Q05066; HTR;
 DR TRANSFAC: T01286;
 DR SGD: S0006269; ROX1;
 DR InterPro: IPR000910; HMG_12_box.
 DR Pfam: PF00505; HMG_box; 1.
 DR SMART: SM00398; HMG; 1.
 KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 FT DNA_BIND 14 83 HMG BOX.
 FT DOMAIN 102 123 GLN-RICH.
 FT SEQUENCE 368 AA; 41838 MW; 3B274442D7DEE3DBD CRC64;
 SO
 Query Match 50.0%; Score 39; DB 1; Length 368;
 Best Local Similarity 33.3%; Pred. No. 32;
 Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 ALNPNTYLSKDVTVG 15
 Db 165 SVNTSNVMSRSLSG 179

RESULT 14
CD36_HUMAN
ID CD36_HUMAN STANDARD: PRT: 471 AA.
PI6671; Q16093; Q13966;
01-AUG-1990 (Rel. 15, Created)
01-AUG-1990 (Rel. 15, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE Platelet glycoprotein IV (GPIV) (GPIIIB) (CD36 antigen) (PAS IV)
(PAS-4 protein).
CD36 OR GP4 OR GP3B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
(1)
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89324065; PubMed=2473841;
RA Oquendo P., Hundt E., Lawler J., Seed B.;
RT "CD36 directly mediates cytoadherence of Plasmodium falciparum
parasitized erythrocytes.";
RL Cell 58:95-101(1989).
(2)
RN SEQUENCE FROM N.A.
RA Sugimoto Y., Tsunuo T.;
RL Submitted (xxx-1992) to the EMBL/GenBank/DBJ databases.
(3)
RN SEQUENCE FROM N.A.
RX MEDLINE=94308159; PubMed=7518447;
RA Armelilla A.L., Vega M.A.;
RT "Structural organization of the gene for human CD36 glycoprotein.";
RL J. Biol. Chem. 269:18985-18991(1994).
(4)
RN SEQUENCE FROM N.A.
RC TISSUE=Platelet;
RX MEDLINE=94082337; PubMed=7505064;
RA Wyler B., Davlet L., Borkiewicz H., Bortel J.C., McGregor J.L.;
RT "Cloning of the cDNA encoding human platelet CD36: comparison to PCR
amplified fragments of monocyte, endothelial and HEL cells.";
RL Thromb. Haemost. 70:500-505(1993).
(5)
RN SEQUENCE FROM N.A.
RX MEDLINE=94040811; PubMed=7593552;
RA Taylor K.T., Tang Y., Sodleski D.A., Lipsky R.H.;
RT "Characterization of two alternatively spliced 5'-untranslated exons
of the human CD36 gene in different cell types.";
RL Gene 133:205-212(1993).
(6)
RN SEQUENCE OF 1-36.
RC TISSUE=Platelet;
RX MEDLINE=89214205; PubMed=2468669;
RA Tandon N.N., Lipsky R.H., Burgess W.H., Jamieson G.A.;
RT "Isolation and characterization of platelet glycoprotein IV (CD36).";
RL J. Biol. Chem. 264:7570-7575(1989).
(7)
RN PALMITOYLATION OF CYS-2, CYS-6, CYS-463 AND CYS-465.
RX MEDLINE=96394284; PubMed=8798390;
RA Tao N., Wagner S.J., Lublin D.M.;
RT "CD36 is palmitoylated on both N- and C-terminal cytoplasmic tails.";
RL J. Biol. Chem. 271:22315-22320(1996).
(8)
RN FUNCTION: SEEMS TO HAVE NUMEROUS POTENTIAL PHYSIOLOGICAL
FUNCTIONS. BINDS TO COLLAGEN, THROMBOSPONDIN, ANIONIC
PHOSPHOLIPIDS AND OXIDIZED LDL. MAY FUNCTION AS A CELL ADHESION
MOLECULE. DIRECTLY MEDIATES CYTOADHERENCE OF PLASMODIUM FALCIPARUM
PARASITIZED ERYTHROCYTES. BINDS LONG CHAIN FATTY ACIDS AND MAY
FUNCTION IN THE TRANSPORT AND/OR AS A REGULATOR OF FATTY ACID
TRANSPORT.
(9)
RN SUBCELLULAR LOCATION: Integral membrane protein.
CC SIMILARITY: BELONGS TO THE CD36 FAMILY.
CC DATABASE: NAME=PROW. NOTE=CD guide CD36 entry;
MMW="http://www.ncbi.nlm.nih.gov/prow/cd/cd36.htm".
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M24795; AAA35534.1;
DR EMBL: M98398; AAS8412.1;
DR EMBL: M98399; AAS8413.1;
DR EMBL: 232770; CAAB3662.1;
DR EMBL: 232754; CAAB3662.1; JOINED.
DR EMBL: 232755; CAAB3662.1; JOINED.
DR EMBL: 232756; CAAB3662.1; JOINED.
DR EMBL: 232757; CAAB3662.1; JOINED.
DR EMBL: 232758; CAAB3662.1; JOINED.
DR EMBL: 232759; CAAB3662.1; JOINED.
DR EMBL: 232760; CAAB3662.1; JOINED.
DR EMBL: 232761; CAAB3662.1; JOINED.
DR EMBL: 232762; CAAB3662.1; JOINED.
DR EMBL: 232763; CAAB3662.1; JOINED.
DR EMBL: 232764; CAAB3662.1; JOINED.
DR EMBL: S67532; AAD13993.1;
DR EMBL: L06850; AAA16068.1;
DR PIR: A30989; A30989.
DR PIR: A33357; A33357.
DR Gene: HGNC:1663; CD36.
DR MIM: 173510;
DR InterPro: IPR002159; CD36.
DR Pfam: PF01130; CD36.1.
KW Glycoprotein; transmembrane; Antigen; Cell adhesion; Transport;
KW Receptor; Lipoprotein; Palmitate.
FT INIT_MET 0
FT DOMAIN 1 6
FT TRANSMEM 7 28
FT DOMAIN 29 438
FT TRANSMEM 439 460
FT DOMAIN 461 471
FT LIPID 2 2
FT LIPID 6 6
FT LIPID 463 463
FT LIPID 465 465
FT DISULFID 242 310
FT DISULFID 271 332
FT DISULFID 312 321
FT CARBOHYD 78 78
FT CARBOHYD 101 101
FT CARBOHYD 133 133
FT CARBOHYD 162 162
FT CARBOHYD 204 204
FT CARBOHYD 219 219
FT CARBOHYD 234 234
FT CARBOHYD 246 246
FT CARBOHYD 320 320
FT CARBOHYD 376 376
FT CONFLICT 43 43
FT CONFLICT 237 237
FT CONFLICT 373 373
SO SEQUENCE 471 AA; 52922 MW; 61B1B5E7D661F161 CRC64;

Query Match 50.08; Score 39; DB 1; Length 471;
Best Local Similarity 50.08; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

OY 3 NPDPV-----LISKDVT 14
DB 305 NPDPVCTEKTISKNT 322
LSGUL_SCHPO

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:47:39 : Search time 4.58824 Seconds
(without alignments)
673.616 Million cell updates/sec

Title: US-09-980-370-1

Sequence: 1 ALNDPNYLISKDVTC 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: SPRENBL_21:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_humani:*
5: sp_invertebrate:*
6: sp_mammali:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_planti:*
11: sp_rodenti:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	1097	2	P72136
2	78	100.0	1223	2	Q92NB5
3	78	100.0	1358	2	P96967
4	78	100.0	1687	2	O9R987
5	78	100.0	1704	2	O31836
6	78	100.0	1706	2	O31836
7	78	100.0	1706	2	O31836
8	78	100.0	1733	2	P72136
9	78	100.0	1733	2	P72136
10	78	100.0	1733	2	O52050
11	78	100.0	1732	2	O51817
12	78	100.0	1732	2	O07442
13	47	60.3	334	5	O45985
14	46	59.0	1579	5	O9BH56
15	45	57.7	312	2	O9K183
16	45	57.7	335	10	Q8W470

17	45	57.7	695	11	O54767	O54767 ratius norv
18	45	57.7	841	10	O9C582	O9C582 arabidopsis
19	44	56.4	315	10	O9EX22	O9EX22 arabidopsis
20	44	56.4	499	10	O3Y334	O3Y334 arabidopsis
21	44	56.4	530	10	O9FH33	O9FH33 arabidopsis
22	44	56.4	630	12	O9EM11	O9EM11 amasacta moo
23	44	56.4	639	16	O9CES3	O9CES3 lactococcus
24	44	56.4	752	17	O9HSA1	O9HSA1 halobacteri
25	44	55.1	1221	8	O9TLM7	O9TLM7 cyanidium c
26	43	55.1	290	10	O49669	O49669 arabidopsis
27	43	55.1	461	10	O9C5K4	O9C5K4 arabidopsis
28	43	55.1	518	5	O9VXC4	O9VXC4 dirosophila
29	43	55.1	925	2	O9VXC4	O9VXC4 dirosophila
30	43	55.1	2877	12	O56078	O56078 peanut bud
31	43	55.1	2878	12	O9DS01	O9DS01 watermelon
32	42	53.8	318	10	O9LMA7	O9LMA7 arabidopsis
33	42	53.8	374	10	O9FX93	O9FX93 arabidopsis
34	42	53.8	500	10	O94D20	O94D20 oryza sativ
35	42	53.8	674	5	O23578	O23578 caenorhabdi
36	42	53.8	1844	5	O23579	O23579 caenorhabdi
37	42	53.8	2348	4	O9UNJ2	O9UNJ2 homo sapien
38	41.5	53.2	2165	17	O8T169	O8T169 methanocarc
39	41	52.6	146	10	O8D391	O8D391 methanocarc
40	41	52.6	256	16	O9C639	O9C639 mus muscula
41	41	52.6	363	16	O9C648	O9C648 mus muscula
42	41	52.6	383	2	O46130	O46130 dirosophila
43	41	52.6	370	5	O9D032	O9D032 dirosophila
44	41	52.6	483	2	O45771	O45771 bacteroides

ALIGNMENTS

RESULT 1

ID	PRELIMINARY:	PRT:	1097 AA.
AC P72136	01-FEB-1997 (TRENDArel. 02, Created)		
DT 01-FEB-1997 (TRENDArel. 02, Last sequence update)			
DT 01-JUN-2003 (TRENDArel. 21, Last annotation update)			
DE TonB-linked adhesin precursor.			
GN TIA.			
OS Porphyromonas gingivalis (Bacteroides gingivalis).			
OC Bacteria: CFM group: Bacteroidetes; Bacteroidales; Porphyromonadaceae;			
OC Porphyromonas.			
OX NCBI_TaxID=837;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=97386416; PubMed=9244265;			
RA Aduse-Opoku J., Stanley J.M., Young K.A., Muir J., Rangarajan M.,			
RA Cutlis M.A.;			
RT "The tla gene of Porphyromonas gingivalis W50: a homologue of the			
RT arginine-specific protease precursor (PrpR) which shares sequence			
RT similarity to TonB-linked receptors.";			
RL J. Bacteriol. 179:4778-4788(1997).			
DR EMBL: Y07618; CAA68897.1;			
DR Interpro: IPR000977; DNM_119ase.			
DR Interpro: IPR001769; Peptidase_C25.			
DR Pfam: PF01364; Peptidase_C25.			
DR PROSITE: PS00597; DNM_LIGASE_A1; UNKNOWN_1.			
KW Signal.			
FT SIGNAL.			
FT SEQUENCE	1097 AA: 118731 MW: 733BA337BA21FBB9 CRC64;		
SO	POTENTIAL		
Query Match	100.0%; Score 78; DB 2; Length 1097;		
Best Local Similarity	100.0%; Pred No. 8.1e-05;		
Matches	15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 ALNDPNYLISKDVTC 15		
DB			
	581 ALNDPNYLISKDVTC 595		

RESULT 2

092NB5 PRELIMINARY: PRT: 1223 AA.
 ID 092NB5
 AC 092NB5
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 130k-HMGD (Fragment).
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales;
 OC NCBI_TaxID=837.
 CX NCBI_TaxID=837.
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RX MEDLINE=99143166; PubMed=9988746;
 RA Shibata Y., Hayakawa M., Takiguchi H., Shirota T., Abiko Y.;
 RT "Determination and characterization of the hemagglutinin-associated
 RT short motifs found in Porphyromonas gingivalis multiple gene
 RT products."
 RL J. Biol. Chem. 274:5012-5020(1999).
 DR EMBL: AB019363; BAA3431.1;
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 1223 AA; 131542 MW; 00225CD2BA9P91D3 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1223;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPVLISKDVYG 15
 DB 707 ALPNPVLISKDVYG 721

RESULT 3

P96967 PRELIMINARY: PRT: 1358 AA.
 ID P96967
 AC P96967
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hemagglutinin.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales;
 OC NCBI_TaxID=837.
 CX NCBI_TaxID=837.
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA Han N., Lepine G., Whitlock J., Wojciechowski L., Prognuske-Fox A.;
 RT "Cloning, sequencing and characterization of hagd, a member of the
 RT harsp multigene family in Porphyromonas gingivalis."
 RT Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 RL EMBL: U68486; ABA49691.1;
 DR InterPro: IPR001777; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 1358 AA; 147102 MW; 47FC40B25006DE08 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPVLISKDVYG 15
 DB 850 ALPNPVLISKDVYG 864

RESULT 4

09R9B7 PRELIMINARY: PRT: 1687 AA.
 ID 09R9B7
 AC 09R9B7
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hemagglutinin/Protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales;
 OC NCBI_TaxID=837.
 CX NCBI_TaxID=837.
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA Han N., Dong H., Prognuske-Fox A.;
 RT "Cloning and characterization of hage from P. gingivalis 381."
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF026946; AAD01810.1;
 DR HSSP: P95493; ICVR.
 DR MEROPS: C25.001;
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT PROTEASE
 SQ SEQUENCE 1687 AA; 183702 MW; D085B516A399E70 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1687;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPVLISKDVYG 15
 DB 1179 ALPNPVLISKDVYG 1193

RESULT 5

051816 PRELIMINARY: PRT: 1704 AA.
 ID 051816
 AC 051816
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Arg-gingipain-1 protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales;
 OC NCBI_TaxID=837.
 CX NCBI_TaxID=837.
 RN 11
 RP SEQUENCE FROM N.A.
 RC MEDLINE=9518080; PubMed=7836351;
 RA Rayvolf N., Potempa J., Pike R.N., Prochazka V., Kiefer M.C.,
 RA Travis J., Barr P.J.;
 RT "Molecular cloning and structural characterization of the Arg-
 RT gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
 RT proteinase-adhesin polypeptide."
 RL J. Biol. Chem. 270:1007-1010(1995).
 DR EMBL: U15282; AAB69539.1;
 DR HSSP: P95493; ICVR.
 DR MEROPS: C25.001;
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT CHAIN 228 719 MATURER-50-KDA CYSTEINE PROTEINASE
 SQ SEQUENCE 1704 AA; 185436 MW; 6A34B40131C2A676 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1704;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
1196 ALNDPNYLISKDVTG 1210

RESULT 6
051839 PRELIMINARY: PRT: 1706 AA.

ID 051839
AC 051839: 051840: PRT: 1706 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
GN PRT.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirszenbaum L., Sotiropoulos C., Jackson C., Cleal S., Slakeski N.,
RA Reynolds E.C.;
RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
gingivalis W50 encoding a 132 kDa protein that contains an arginine-
specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Blochem. Biophys. Res. Commun. 207:424-431(1995).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=96311339; PubMed=8713096;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes
an arginine-specific thiol proteinase and multiple adhesins.";
RL Blochem. Biophys. Res. Commun. 224:605-610(1996).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Reynolds E.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN 14
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Slakeski N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L26341; AAC18876.1; -.
DR HSSP; P95493; 1CVR.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
GN KGP.
FT SIGNAL.
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 228 1706 ARGININE-SPECIFIC THIOL PROTEASE.
SQ SEQUENCE 1706 AA; 185626 MW; E8BDFO7C9813844 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
1198 ALNDPNYLISKDVTG 1212

RESULT 7
051838 PRELIMINARY: PRT: 1706 AA.
ID 051838
AC 051838;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Protease precursor.
GN PRT.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=96071894; PubMed=7591131;
RA Aduse-Opoku J., Muir J., Slaney J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease
antigen (PRP1) of Porphyromonas gingivalis W50.";
RL Infect. Immun. 63:4744-4754(1995).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Rangarajan M., Aduse-Opoku J., Slaney J.M., Young K.A., Curtis M.A.;
RT "The prp1 and the pr2 arginine-specific protease genes of
Porphyromonas gingivalis W50 produce five biochemically distinct
enzymes.";
RL Mol. Microbiol. 23:0-0(1997).
DR EMBL; X82680; CAA57997.1; -.
DR HSSP; P95493; 1CVR.
DR MEROPS; C25.001; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
GN SIGNAL.
FT SIGNAL.
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E56DCD87FDA8CDD CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
1198 ALNDPNYLISKDVTG 1212

RESULT 8
P72194 PRELIMINARY: PRT: 1723 AA.
ID P72194
AC P72194;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
FT SIGNAL.
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E56DCD87FDA8CDD CRC64;
Query Match 100.0%; Score 78; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1723 AA; 187261 MW; 5628963D51493EB CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 1215 ALNPDNYLISKDVTG 1229

RESULT 9

P72197 PRELIMINARY; PRT: 1723 AA.

AC 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lys-gingipain.
 GN KGP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NC NCB1_TaxID=837;
 RN (1)
 RA SEQUENCE FROM N.A.
 RP Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
 RA Prochazka V., Klefer M.C., Travis J., Barr P.J.;
 RC "Molecular cloning and characterization of Porphyromonas gingivalis
 RT Lys-gingipain.";
 RT Submitted (Apr. 1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U04691; AAA9810.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR00977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1723 AA; 186831 MW; 4508A7E50197CEBD CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 1215 ALNPDNYLISKDVTG 1229

RESULT 10

O52050 PRELIMINARY; PRT: 1732 AA.

DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine specific cysteine protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NC NCB1_TaxID=837;
 RN (1)
 RA SEQUENCE FROM N.A.
 RP STRAIN-W83.
 RX MEDLINE-98298016; PubMed-9632563;
 RA Lewis J.P., Macrina F.L.;
 RT "IS195, an insertion sequence-like element associated with protease
 genes in Porphyromonas gingivalis.";
 RL Infect. Immun. 66:3035-3042(1998).
 DR EMBL: AF017059; AAC26523.1; -.
 DR MEROPS: C25.002; -.

DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA; 187931 MW; 82337463D5CB5EA5 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 1216 ALNPDNYLISKDVTG 1230

RESULT 11

O51817 PRELIMINARY; PRT: 1732 AA.

AC 051817;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Porphylin.
 GN PRP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NC NCB1_TaxID=837;
 RN (1)
 RA SEQUENCE FROM N.A.
 RP STRAIN-W12;
 RX MEDLINE-96213011; PubMed-8631659;
 RA Barkocy-Gallagher G.A., Han N., Patti J.M., Whitlock J.,
 RA Progulski-Fox A., Lantz M.S.;
 RC "Analysis of the prtP gene encoding porphylin, a cysteine proteinase
 RT of Porphyromonas gingivalis.";
 RT J. Bacteriol. 178:2734-2741(1996).
 DR EMBL: U42210; AAB05655.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA; 187875 MW; 654271DBE7BCAE4 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 1216 ALNPDNYLISKDVTG 1230

RESULT 12

O07442 PRELIMINARY; PRT: 1732 AA.

AC 007442;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine-specific cysteine proteinase.
 GN PRP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NC NCB1_TaxID=837;
 RN (1)
 RA SEQUENCE FROM N.A.
 RP STRAIN-W50;
 RX MEDLINE-99235907; PubMed-10219167;
 RA Stakeski N., Cleal S.M., Bhogal P.S., Reynolds E.C.;

RT *Characterization of a Porphyromonas gingivalis gene prtx that encodes
RT a lysine-specific cysteine proteinase and three sequence-related

RL Oral Microbiol. Immunol. 14:92-97(1999).

DR EMBL; U75366; AAB50809.1; -.

DR MEMOPS; C25.002; -.

DR InterPro; IPR000977; DNA_Ligase.

DR InterPro; IPR001769; Peptidase_C25.

DR Pfam; PF01364; Peptidase_C25; 3.

DR PROSITE; PS00697; DNA_LIGASE.A1; UNKNOWN.1.

SO SEQUENCE 1732 AA; 167914 MW; 45D5B91377391703 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1732;

Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDYTG 15

DB 1216 ALNPDNYLSKDYTG 1230

RESULT 13

045985

AC 045985; PRELIMINARY; PRT; 334 AA.

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 2K1037.3 protein.

GN 2K1037.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OX NCBI_TaxID=6239;

RA none;

RP SEQUENCE FROM N.A.

RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RL Investigating biology "1";

RT Science 282:2012-2018(1998).

DR EMBL; Z81142; CA03502.1;

DR InterPro; IPR002851; DUF32.

DR Pfam; PF01748; DUF32; 1.

SO SEQUENCE 334 AA; 37719 MW; E8592D5A4D24FBA3 CRC64;

Query Match 60.3%; Score 47; DB 5; Length 334;

Best Local Similarity 60.2%; Pred. No. 5.9;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LNPDNYLSKDYT 14

DB 22 LNPDNYLSEEDVT 34

RESULT 14

09BH56

AC 09BH56; PRELIMINARY; PRT; 1579 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE M01G12.12 protein.

GN M01G12.12.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OX NCBI_TaxID=6239;

RA none;

RP SEQUENCE FROM N.A.

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery."

DR EMBL; AF237553; ABE81413.1;

DR InterPro; IPR001769; Peptidase_C25.

DR Pfam; PF01364; Peptidase_C25; 2.

SO SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25P5 CRC64;

RP SEQUENCE FROM N.A.

RA McJury K.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RL MEDLINE=99069613; PubMed=9851916;

RA none;

RP "Genome sequence of the nematode C. elegans: A platform for

RT investigating biology."

RT Science 282:2012-2018(1998).

RL 131

RP SEQUENCE FROM N.A.

RA Lennard N.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL132848; CAC35915.1; -.

DR EMBL; Z81571; CAC35915.1; JOINED.

DR EMBL; Z81571; CAC35914.1; -.

DR EMBL; AL132848; CAC35914.1; JOINED.

SO SEQUENCE 1579 AA; 182870 MW; B645172007155E26C CRC64;

Query Match 59.0%; Score 46; DB 5; Length 1579;

Best Local Similarity 46.7%; Pred. No. 47;

Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDYTG 15

DB 601 SLGPDNYLTHDLG 615

RESULT 15

09KIB3

AC 09KIB3; PRELIMINARY; PRT; 312 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical outer membrane protein PG27.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;

OX NCBI_TaxID=837;

RA none;

RP SEQUENCE FROM N.A.

RL STRAIN-W50;

RA Ross B., Barr I., Patterson M., Agius C., Rothel L., Margelits M.,

RT "P. gingivalis polypeptides and nucleic acids."

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RL STRAIN-W50;

RA Ross B.C.;

RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery."

DR EMBL; AF237553; ABE81413.1;

DR InterPro; IPR001769; Peptidase_C25.

DR Pfam; PF01364; Peptidase_C25; 2.

SO SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25P5 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 312;

Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDY 13

DB 53 ALNPDNYLTPEL 65

Search completed: June 26, 2003, 00:53:25
Job time: 7.58824 secs

Thu Jun 26 11:58:08 2003

us-09-980-370-5.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 26, 2003, 00:36:49 ; Search time 31 Seconds
(without alignments)
763,098 Million cell updates/sec

Title: US-09-980-370-5
Sequence: 1 gcaagctcagcgaagcgtc.....ccaggaagcgaatgagcgc 402

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Command line parameters:

-MODE=frame-n2p model -DEV=xxh
-O=/cgn2.1/USPTO.spool/US09980370.runal.19062003.17443.8879/app.query.fasta.1.583
-DB=Issued.Patents.AA -OPWT=fastan -SUPRT=rai -MINMATCH=0.1 -LOOEXT=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humand40.cdl
-LIST=45 -DOALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980370.gcg.1.1.13.gunal.19062003.17443.8879 -NCPU=6 -ICPU=3
-NO_MMAP -LANG=ENGLISH -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-NO_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued.Patents.AA.*
1: /cgn2.6/ptodata/1/aa/5A.COMB.pep.*
2: /cgn2.6/ptodata/1/aa/5B.COMB.pep.*
3: /cgn2.6/ptodata/1/aa/6A.COMB.pep.*
4: /cgn2.6/ptodata/1/aa/6B.COMB.pep.*
5: /cgn2.6/ptodata/1/aa/PTUS.COMB.pep.*
6: /cgn2.6/ptodata/1/aa/Beckfilesl.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result.	Query	Match	Length	DB	ID	Description
No.						
1	716	87.9	1687	2	US-08-570-311-29	Sequence 29, Appl
2	716	87.9	1704	3	US-08-336-308A-10	Sequence 10, Appl
3	716	87.9	1704	3	US-08-822-324-6	Sequence 6, Appl
4	716	87.9	1704	4	US-09-490-931-10	Sequence 10, Appl
5	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
6	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
7	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
8	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
9	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
10	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
11	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl
12	716	87.9	1704	4	US-08-570-311-10	Sequence 10, Appl

13	704	96.3	456	2	US-08-570-311-18	Sequence 18, Appl
14	704	96.3	456	2	US-08-570-311-20	Sequence 20, Appl
15	176.5	24.1	497	2	US-08-570-311-2	Sequence 2, Appl
16	176.5	24.1	497	2	US-08-353-465-2	Sequence 2, Appl
17	83.5	11.4	316	3	US-08-682-643-4	Sequence 4, Appl
18	79.5	10.9	316	1	US-08-038-932B-1	Sequence 1, Appl
19	79.5	10.9	316	1	US-08-656-349-1	Sequence 1, Appl
20	79.5	10.9	316	4	US-09-104-623A-5	Sequence 5, Appl
21	79.5	10.9	316	4	US-09-019-533-5	Sequence 5, Appl
22	79.5	10.9	316	4	US-09-285-108-2	Sequence 2, Appl
23	78	10.7	785	3	US-09-479-264-2	Sequence 2, Appl
24	75.5	10.8	2972	4	US-09-579-181-2	Sequence 2, Appl
25	75.5	10.8	3118	4	US-09-579-181-1	Sequence 1, Appl
26	73.5	10.1	319	3	US-08-682-643-3	Sequence 3, Appl
27	73	10.0	435	4	US-09-072-917A-9	Sequence 9, Appl
28	72.5	9.9	265	3	US-08-483-857-8	Sequence 8, Appl
29	72.5	9.9	617	4	US-09-188-930-303	Sequence 303, App
30	71.5	9.8	1422	4	US-09-134-001C-4493	Sequence 4493, Ap
31	71.5	9.8	1422	4	US-08-459-260A-82	Sequence 82, Appl
32	71	9.7	236	3	US-08-886-269-2	Sequence 2, Appl
33	71	9.7	236	3	US-08-886-269-2	Sequence 2, Appl
34	71	9.7	236	3	US-08-886-269-2	Sequence 2, Appl
35	69.5	9.5	2123	4	US-08-560-398-2	Sequence 10, Appl
36	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
37	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
38	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
39	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
40	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
41	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
42	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
43	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
44	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl
45	69	9.4	271	3	US-08-968-685A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-570-311-29
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiskee-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ted W. WhiteLock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09/05/94
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119

Thu Jun 26 11:58:08 2003

us-09-980-370-5.ra1

Page 2

```

: FILING DATE: 25-JAN-1991
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA: US 07/241,640
: APPLICATION NUMBER: 08-SEP-1988
: FILING DATE: 08-SEP-1988
: ATTORNEY/AGENT INFORMATION:
: NAME (MILLER, Ted M.)
: REGISTRATION NUMBER: 36,965
: REFERENCE/DOC. NUMBER: 015-C3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (904) 372-5800
: TELEFAX: (904) 372-5800
: INFORMATION FOR SEQ ID NO: 29:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1687 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-570-311-29

Alignment Scores:
Pred. No.: 4,68e-77 Length: 1687
Score: 716.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.95% Indels: 0
Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-29 (1-1687)
OY 1 GCAGACTTCACGGAACGCTTCAGCTCTTCTACTCATGAGAGGACACGAGCAATGAGCT 60
DB 1120 AlaaSpheThrChLurhPhelGluSerSerThrHisGlyAlaProAlaGluTrpThr 1139
OY 61 ACTATCGATGCCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
DB 1140 ThrLleSpAlaSpGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 1159
OY 121 TGCTTCACAGCTCAAGCGGCGACACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 180
DB 1160 TrpLeuThrAlaHisGlyGlyTrpHisValAlaLleSerPheSerTrpAsnGlyMetAla 1179
OY 181 TTCAATCGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 240
DB 1180 LeuAsnProAspAsnTrpLeuLleSerLysAspValThrGlyAlaThrLysValLysTr 1199
OY 241 TACTATCGATCAAGCGGCTTCCGCGGATGAGTACGAGTACGAGTACGAGTACGAGTAC 300
DB 1200 TyrTrpAlaValAsnAspGlyPheProGlyAspHisTrpAlaValMetLleSerLysTr 1219
OY 301 GCGAGAACGCGGAGACTTCAGCGTGTTCGAGAACGCGGAGAACGCGGAGAACGCGGAG 360
DB 1220 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluTrpProAsnGlyLleAsnLys 1239
OY 361 GCGGAGAACGATTCGCTCTTCACGGAACGCAATGCGGCG 402
DB 1240 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1253

RESULT 2
US-08-336-308A-10
: Sequence 10, Application US/08336308A
: GENERAL INFORMATION:
: APPLICANT: Travis, James
: APPLICANT: Potempa, Jan S.
: APPLICANT: Bart, Philip J.
: APPLICANT: Pavloff, Nadine
: TITLE OF INVENTION: Porphyromonas gingivalis
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS: Arginine-specific Proteinase Coding Sequences
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Circle, Suite 201
```

```

: CITY: Boulder
: STATE: Colorado
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/336,308A
: FILING DATE: 08-NOV-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/119,361
: FILING DATE: 10-SEP-1993
: PRIOR APPLICATION DATA: US 08/265,441
: FILING DATE: 24-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Fester, Donna M., 33, 878
: REGISTRATION NUMBER: 33,878
: REFERENCE/DOC. NUMBER: 33,878
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 449-8080
: TELEFAX: (303) 449-8089
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1704 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-336-308A-10

Alignment Scores:
Pred. No.: 4,68e-77 Length: 1704
Score: 716.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.95% Indels: 0
Gaps: 0

US-09-980-370-5 (1-402) x US-08-336-308A-10 (1-1704)
OY 1 GCAGACTTCACGGAACGCTTCAGCTCTTCTACTCATGAGAGGACACGAGCAATGAGCT 60
DB 1137 AlaaSpheThrChLurhPhelGluSerSerThrHisGlyAlaProAlaGluTrpThr 1156
OY 61 ACTATCGATGCCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
DB 1157 ThrLleSpAlaSpGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 1176
OY 121 TGCTTCACAGCTCAAGCGGCGACACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 180
DB 1177 TrpLeuThrAlaHisGlyGlyTrpHisValAlaLleSerPheSerTrpAsnGlyMetAla 1196
OY 181 TTCAATCGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 240
DB 1197 LeuAsnProAspAsnTrpLeuLleSerLysAspValThrGlyAlaThrLysValLysTr 1216
OY 241 TACTATCGATCAAGCGGCTTCCGCGGATGAGTACGAGTACGAGTACGAGTACGAGTAC 300
DB 1217 TyrTrpAlaValAsnAspGlyPheProGlyAspHisTrpAlaValMetLleSerLysTr 1236
OY 301 GCGAGAACGCGGAGACTTCAGCGTGTTCGAGAACGCGGAGAACGCGGAGAACGCGGAG 360
DB 1237 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluTrpProAsnGlyLleAsnLys 1256
OY 361 GCGGAGAACGATTCGCTCTTCACGGAACGCAATGCGGCG 402
DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1270

RESULT 3
```



```

OY 61 ACTATGATGCGATGCGATGAGGTTGGCTCTGCTCTTCCGCAATATGGAC 120
DB 1157 ThrileaspAlaaspGlyAspGlyGlnGlyTrpleucyLeuSerSerGlyGlnLeuasp 1176
OY 121 TGGCTCACAGCTGATGCGGACCAACAGTAGTACGCTTTCTGATCGAATGAAATGGCT 180
DB 1177 TrpleuthrAlaHisGlyGlyThrAsnValValAlaIsePheSerTrpAsnGlyMetAla 1196
OY 181 TTGAATCCTGATACATATCTCATCAAGGATGTTACAGCGCAACAGTAAGTAC 240
DB 1197 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1216
OY 241 TACTATCCAGTCAACGAGCGTTTCCCGGATCACTATGCGGATGATCTCCAAGC 300
DB 1217 TTTTAAAlaValAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1236
OY 301 GCGACGACGCGCGAGACTTCAGCGTTGTTTCGAGAAAGCGCTAACGGAATAATAG 360
DB 1237 GlyThrAsnAlaGlyAspPheThrValValAlaPheGlnGlnThrProAsnGlyIleAsnLys 1256
OY 361 GCGGAGCAAGATTGCGTCTTTCACAGGAAAGCAATGCGGCC 402
DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGlnAlaAsnGlyAla 1270

RESULT 5
US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Leptine, Guyalaine
APPLICANT: Han, Naimling
APPLICANT: Lantcz, Marilyn
TITLE OF INVENTION: Cloned Porphyromonas gingivallis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570.311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353.485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647.119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241.640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: 0815.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10

Alignment Scores:
Pred. No.: 6,2e-77 Length: 1732
Score: 715.00 Matches: 131
Percent Similarity: 99.254 Conservative: 2
Best Local Similarity: 97.764 Mismatches: 1
Query Match: 97.814 Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-10 (1-1732)
OY 1 GCGACTTCACGGAACGTTGAGTCTTCTACTGATGAGAGCACCAGCAATGAGCT 60
DB 1157 ThrileaspAlaaspGlyAspGlyGlnGlyTrpleucyLeuSerSerGlyGlnLeuasp 1176
OY 61 ACTATGATGCGATGCGATGAGGTTGGCTCTGCTCTTCCGCAATATGGAC 120
DB 1177 ThrileaspAlaaspGlyAspGlyGlnGlyTrpleucyLeuSerSerGlyGlnLeuasp 1196
OY 121 TGGCTCACAGCTGATGCGGACCAACAGTAGTACGCTTTCTGATCGAATGAAATGGCT 180
DB 1197 TrpleuthrAlaHisGlyGlyThrAsnValValAlaIsePheSerTrpAsnGlyMetAla 1216
OY 181 TTGAATCCTGATACATATCTCATCAAGGATGTTACAGCGCAACAGTAAGTAC 240
DB 1217 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1236
OY 241 TACTATCCAGTCAACGAGCGTTTCCCGGATCACTATGCGGATGATCTCCAAGC 300
DB 1237 TTTTAAAlaValAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1256
OY 301 GCGACGACGCGCGAGACTTCAGCGTTGTTTCGAGAAAGCGCTAACGGAATAATAG 360
DB 1257 GlyThrAsnAlaGlyAspPheThrValValAlaPheGlnGlnThrProAsnGlyIleAsnLys 1276
OY 361 GCGGAGCAAGATTGCGTCTTTCACAGGAAAGCAATGCGGCC 402
DB 1277 GlyGlyAlaArgPheGlyLeuSerThrGlnAlaAsnGlyAla 1290

RESULT 6
US-08-353-485-10
Sequence 10, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Leptine, Guyalaine
APPLICANT: Han, Naimling
APPLICANT: Lantcz, Marilyn
TITLE OF INVENTION: Cloned Porphyromonas gingivallis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353.485

```


NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Alignment Scores:
Pred. No.: 2,18e-76 Length: 2628
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-14 (1-2628)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGCACCAGCGAATGAGCT 60
Db 685 AlaaspphetrhrgluthrphggluserSerThrhlsGlyAlaProAlaGluTrpThr 704
QY 61 ACATATGATCCGATGCGGATGGTGGAGGCTCTGCTGCTCTTCCGACATTTGAC 120
Db 705 ThrllasprAlaasprGlyaspGlynglYtrleucysleuSerSerGlyGlnLeuasp 724
QY 121 TGCGTCACAGCTCATGCGCGCACCAACGTAGTAACTCTTTCATGATGAATGAGCT 180
Db 725 TrpleuthrAlahlsGlyglYthrasnValAlaAserrheserTrpansGlyMetAla 744
QY 181 TTGAATCCGTATACATCTCATCTCAAGAGATGTACAGCGCACGAAGTAAGTAC 240
Db 745 LeuasnpProaspaSnYtrleuIlleSerlyAspValThrGlyAlaThrLysVallystyr 764
QY 241 TACTATCCAGTCACACGCGTTTCCCGGGATCAGTACATGCGGTGATGATCTCCAACAG 300
Db 765 TytrYtrAlaValasnaSpGlyaspheProGlyAspHstYtrAlaValMetIlleSerLysThr 784
QY 301 GGCAGCAAGCCGCGAGACTTCACGCGTTGTTTCGAAGAAGCCCTAAGCAATGAATAG 360
Db 785 GlyThrAsnaIaaspGlyaspheThrValAlaPhgglucluthrProasnglylIleasnLys 804
QY 361 GCGGAGCAAGATTGCGTCTTTCACGGAAGCCCAATGGCGCC 402
Db 805 GlyGlyAlaIarGphgglYleuSerThrGluAlaaspGlyAla 818

RESULT 10
US-08-570-311-8
Sequence 8, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proquilske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiping
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570.311
FILING DATE:
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-8

Alignment Scores:
Pred. No.: 2,71e-76 Length: 1087
Score: 709.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.99% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-8 (1-1087)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGCACCAGCGAATGAGCT 60
Db 520 AlaaspphetrhrgluthrphggluserSerThrhlsGlyAlaProAlaGluTrpThr 539
QY 61 ACATATGATCCGATGCGGATGGTGGAGGCTCTGCTGCTCTTCCGACAAATGGAC 120
Db 540 ThrllasprAlaasprGlyaspGlynglYtrleucysleuSerSerGlyGlnLeuasp 559
QY 121 TGCGTCACAGCTCATGCGCGCACCAACGTAGTAACTCTTTCATGATGAATGAGCT 180
Db 560 TrpleuthrAlahlsGlyglYthrasnValAlaAserrheserTrpansGlyMetAla 579
QY 181 TTGAATCCGTATACATCTCATCTCAAGAGATGTACAGCGCACGAAGTAAGTAC 240
Db 580 LeuasnpProaspaSnYtrleuIlleSerlyAspValThrGlyAlaThrLysVallystyr 599
QY 241 TACTATCCAGTCACACGCGTTTCCCGGGATCAGTACATGCGGTGATGATCTCCAACAG 300
Db 600 TytrYtrAlaValasnaSpGlyaspheProGlyAspHstYtrAlaValMetIlleSerLysThr 619
QY 301 GGCAGCAAGCCGCGAGACTTCACGCGTTGTTTCGAAGAAGCCCTAAGCAATGAATAG 360
Db 620 GlyThrAsnaIaaspGlyaspheThrValAlaPhgglucluthrProasnglylIleasnLys 639
QY 361 GCGGAGCAAGATTGCGTCTTTCACGGAAGCCCAATGGCGCC 402
Db 640 GlyGlyAlaIarGphgglYleuSerThrGluAlaaspGlyAla 653

RESULT 11
US-08-353-485-8
; Sequence 8, Application US/08353485
; Patent No. 5830710
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tummasorn, Somying
; APPLICANT: Lepine, Guyalaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalls Genes
; TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1087 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-353-485-8

Alignment Scores:
Pred. No.: 2,71e-76 Length: 1087
Score: 709.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.99% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-353-485-8 (1-1087)

OY 1 GCAGACTTCAGCGAAGCTTCGATCTCATGAGAGGACGACGAGGATGACT 60
DB 520 AIAAspRhetnrgLutHrPheGluSerSerThrnHsGlyGluAlaProAlaGluLutPrHr 539
OY 61 ACTATCGATCGCGATGGATGGGCTCGTCCTCTCCGAGCAATTCGAC 120
DB 540 ThrIleAspAlaAspGlyAspGlyGlnAspTrpLeuCySleuSerSerGlyGlnLeuAsp 559
OY 121 TGGCTCAGACGCTCATGCGCGGACGACGACGATAGTACGCTTTCTCATGGAATGGAGCT 180
DB 560 TrpLeuThrAlaHnHsGlyGlyThrnAsnValValAlaLaserPheSerTrpAsnGlyMetAla 579

OY 181 TTGAATCCTGATTAATCATCTCATCAAGGATTTACAGGCGCAAGGTAAGTAC 240
DB 580 LeuAsnProAspsnrryLeuIleSerLysAspValThrGlyLalHrLysValLysTyr 599
OY 241 TACTATCAGTCAAGCAGGTTTCCCGGGATCAGTACGCGGTGATGATCCAGACG 300
DB 600 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 619
OY 301 GCGCAGACGCGCGAGCTTCACGCTGTTTCGAGAAACGCTAACGGAATAAATPAG 360
DB 620 GlyThrnAsnAlaGlyAspHetrnValAlaPheGluGluLutHrProAsnGlyIleAsnLys 639
OY 361 GCGGAGCAAGATTTCGTTTCACGAGGAAGCCCAATGGCGCC 402
DB 640 GlyGlyAlaArpHegLysSerTrnGluAlaAsnGlyAla 653

RESULT 12
US-08-570-311-27
; Sequence 27, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tummasorn, Somying
; APPLICANT: Lepine, Guyalaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalls Genes
; TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1358 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-570-311-27

Alignment Scores:

Pred. No.:	2,95e-76	Length:	1358
Score:	709.00	Matches:	130
Percent Similarity:	99.51%	Conservative:	2
Best Local Similarity:	97.01%	Mismatches:	2
Query Match:	96.99%	Indels:	0
DB:	2	Gaps:	0

US-09-980-370-5 (1-402) x US-08-570-311-27 (1-1358)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGGACGACGGAATGACT 60
 DB 791 AlAspPheThrGluThrPheGluSerSerThrHisGlyAlaIleProAlaGluTrpThr 810
 QY 61 ACTATGATCCGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 120
 DB 811 ThrIleAspAlaIleAspGlyAspGlyAlaIleAspGlyAlaIleAspGlyAlaIleAsp 830
 QY 121 TGGCTCAGACGTCATGAGGAGGACGACGACGACGACGACGACGACGACGACGACGACG 180
 DB 831 TrpLeuThrAlaHisGlyGlyThrAsnValAlaIleAspPheSerThrPheAsnGlyMetAla 850
 QY 181 TTGAATCCTGATACATATCTCATCTCAAGAGATGTTACAGGACGACGACGACGACGACGAC 240
 DB 851 LeuAsnProAspAsnIleuIleSerIleuIleSerIleuIleSerIleuIleSerIleuIle 870
 QY 241 TACTATCCAGTCAACGACGCTTTCCGGGATCAGTATGCGGATGATGATGATGATGATGAT 300
 DB 871 TyrTrpAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerIleThr 890
 QY 301 GGCAGACGACGCGGACGACGCTTTCGAGAGAACGCTTAACGAAATAAATAG 360
 DB 891 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnIle 910
 QY 361 GCGGAGCAAGATTGCTTTCACGAGGACGACGACGACGACGACGACGACGACGACGACG 402
 DB 911 GlyGlyAlaIleArgPheGlyLeuSerThrGluAlaAsnGlyAla 924

RESULT 13
 US-08-570-311-18
 Sequence 18, Application US/08570311
 Patent No. 5824791
 GENERAL INFORMATION:
 APPLICANT: Proguis-Ske-Fox, Ann
 APPLICANT: Tumwasorn, Somying
 APPLICANT: Lepline, Guyelaine
 APPLICANT: Han, Naiming
 APPLICANT: Lantz, Marilyn
 TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ted W. Whitlock
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/570.311
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/353,485
 FILING DATE: 09-DEC-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119
 FILING DATE: 25-JAN-1991
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/241,640
 FILING DATE: 08-SEP-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Whitlock, Ted W.
 REGISTRATION NUMBER: 36,965
 REFERENCE/DOCKET NUMBER: UF15.C3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904) 375-8100
 TELEFAX: (904) 372-5800
 INFORMATION FOR SEQ. ID NO.: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 456 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-570-311-18

Alignment Scores:

Pred. No.:	7,74e-76	Length:	456
Score:	704.00	Matches:	129
Percent Similarity:	98.51%	Conservative:	3
Best Local Similarity:	96.27%	Mismatches:	2
Query Match:	96.31%	Indels:	0
DB:	2	Gaps:	0

US-09-980-370-5 (1-402) x US-08-570-311-18 (1-456)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGGACGACGGAATGACT 60
 DB 192 AlAspPheThrGluThrPheGluSerSerThrHisGlyAlaIleProAlaGluTrpThr 211
 QY 61 ACTATGATCCGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 120
 DB 212 ThrIleAspAlaIleAspGlyAspGlyAlaIleAspGlyAlaIleAspGlyAlaIleAsp 231
 QY 121 TGGCTCAGACGTCATGAGGAGGACGACGACGACGACGACGACGACGACGACGACGACG 180
 DB 232 TrpLeuThrAlaHisGlyGlyThrAsnValAlaIleAspPheSerThrPheAsnGlyMetAla 251
 QY 181 TTGAATCCTGATACATATCTCATCTCAAGAGATGTTACAGGACGACGACGACGACGACGAC 240
 DB 252 LeuAsnProAspAsnIleuIleSerIleuIleSerIleuIleSerIleuIleSerIleuIle 271
 QY 241 TACTATCCAGTCAACGACGCTTTCCGGGATCAGTATGCGGATGATGATGATGATGATGAT 300
 DB 272 TyrTrpAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerIleThr 291
 QY 301 GGCAGACGACGCGGACGACGCTTTCGAGAGAACGCTTAACGAAATAAATAG 360
 DB 292 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnIle 311
 QY 361 GCGGAGCAAGATTGCTTTCACGAGGACGACGACGACGACGACGACGACGACGACGACG 402
 DB 312 GlyGlyAlaIleArgPheGlyLeuSerThrGluAlaAspGlyAla 325

RESULT 14
 US-08-570-311-20
 Sequence 20, Application US/08570311
 Patent No. 5824791
 GENERAL INFORMATION:
 APPLICANT: Proguis-Ske-Fox, Ann
 APPLICANT: Tumwasorn, Somying
 APPLICANT: Lepline, Guyelaine
 APPLICANT: Han, Naiming
 APPLICANT: Lantz, Marilyn
 TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
 NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 372-5800
TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-570-311-20

Alignment Scores:

pred. No.: 7.74e-76 Length: 456
Score: 704.00 Matches: 129
Percent Similarity: 58.51% Conservative: 2
Best Local Similarity: 96.27% Mismatches: 0
Query Match: 96.31% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-20 (1-456)

QY 1 GCAGACTCAGGAAACCTTGAGCTCTTCTATGAGAGGACGACGAGGAATGACT 60
DB 192 AAlaAspPheThrGluThrPheGluSerSerThrHisGlyLysAlaProAlaGluTyrThr 211
QY 61 ACTATGATGCGGATGCGATGATGAGGAGTGGCTGTCTGTCTTCCGACAATTGAC 120
DB 212 ThrIleAspAlaAspGlyAspGlyGlnGlyTyrPheGluSerSerGlyGlnGly 231
QY 121 TGGCTCAGACTGATGAGGCGACCACTACTAAGCTTCTCATGGAATGGAATGCT 180
DB 232 TrpLeuThrAlaHisGlyGlyThrAsnValAlaLaserPheSerTrpAsnGlyMetAla 251
QY 161 TTGAATCTGATTAACATCTCATCTCAAGAGATGTAAGCGACGACGAAGTAAGTAC 240
DB 252 LeuAsnProAspAspTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 271
QY 241 TACTATCAGTCAAGCAAGAGGTTTCCGCGGATGATGATGATGATGATGATGATGAT 300
DB 272 TyrTyrAlaValAlaAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 291
QY 301 GGCACGAAGCGCGGAGACTTACGCGTGTGTTTCAGAGAAGACGCTTAAGCAATTAATAG 360

DB 292 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 311

QY 361 GCGGAGCAAGATTCGCTCTTCCAGCGAGCAATGCGCC 402
DB 312 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAspGlyAla 325

RESULT 15

US-08-570-311-2
Sequence 2, Application US/08570311
Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Maiming
APPLICANT: Lantz, Marilyn
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 372-5800
TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-570-311-2

Alignment Scores:

pred. No.: 9.55e-13 Length: 497
Score: 176.50 Matches: 45
Percent Similarity: 51.69% Conservative: 16
Best Local Similarity: 38.14% Mismatches: 44
Query Match: 24.15% Indels: 13
DB: 2 Gaps: 5

US-09-980-370-5 (1-402) x US-08-570-311-2 (1-497)

```

OY 13 GAAACGTTGACGCTCTTCTACTCATGGAGAGCAGCGAATGCACTACTATCGATGCC 72
    |||:|||||:||||| ||| ||||| |||||
Db 235 GluSerPheAspThrGlnThr-----LeuProAsnGlyTrpThrMetIleAspAla 251
    |||:|||||:||||| ||| ||||| |||||
OY 73 GATGGCGATGGTGAAGGTTGGCTGTCTGTCTTCGACAAATGGACTGCTCAGACT 132
    |||:|||||:||||| ||| ||||| |||||
Db 252 AspGlyAspGlyHisAsnTrp-----LeuSerThrIleAsnValTyrAsnThrAlaThr 269
    |||:|||||:||||| ||| ||||| |||||
OY 133 CATGGCGGCGACCAACGTAAGCTCTTCTCATGGATGGA-----ATG 177
    ||| ||| ||| ||| |||||
Db 270 HisThrClyAspGlyAlaMetPheSerLysSerTrpThrAlaSerClyGlyAlaLysIle 289
    ||| ||| ||| ||| |||||
OY 178 GCTTGAATCCTGATACTATCTCATCTCAAGATGTTACAGCGCGCAAGAGGTAAG 237
    |||:|||||:||||| ||| ||||| |||||
Db 290 AspLeuSerProAspAsnTyrIleuValThrProLysValThrValProGluAsnGlyLys 309
    |||:|||||:||||| ||| ||||| |||||
OY 238 TACTACTATCCAGTCAACGACGAGGTTTCCC-----GGGATCAGTATGCGGTGATGATC 291
    ||| ||| ||| ||| |||||
Db 310 LeuSerTyrTrpValSerSerGlnValProTrpThrAsnGluHisTyrClyValPheLeu 329
    |||:|||||:||||| ||| ||||| |||||
OY 292 TCCAGAGCGGCGACGAGCGCGAGACTTCACGGT---GTTTCGAGAGAAGC 342
    ||| ||||| ||||| ||||| |||||
Db 330 SerThrThrGlyAsnGluAlaAlaAsnPheThrIleLysLeuLeuGluGluThr 347
    ||| ||||| ||||| ||||| |||||

```

Search completed: June 26, 2003, 00:47:33
 Job time : 36 secs

Page 2

US-09-980-370-5 (1-402) X US-10-146-473-50 (1-2971)

RESULT 2
116-08-738-636-4263

Sequence 4263	Application: US/09738626
Publication No	US20020197005A1
GENERAL INFORMATION	
APPLICANT:	NAKAGAWA, SATOSHI
APPLICANT:	MIYOCUCHI, HIROSHI
APPLICANT:	ANDO, SEIJO
APPLICANT:	HAYASHI, MIKIRO
APPLICANT:	OGHAI, KEIJO
APPLICANT:	YOKOI, HARUHIKO
APPLICANT:	TATEISHI, MOMO
APPLICANT:	SENOM, AKIHITO
APPLICANT:	IKEDA, MASATO
APPLICANT:	OZAKI, AKIO
TITLE OF INVENTION:	NOVEL POLYNUCLEOTIDES
FILE REFERENCE:	249-125
CURRENT APPLICATION NUMBER:	US/09738, 626
CURRENT FILING DATE:	2000-12-18
PRIOR APPLICATION NUMBER:	JP 99/377484
PRIOR FILING DATE:	1999-12-16
PRIOR APPLICATION NUMBER:	JP 00/159162
PRIOR FILING DATE:	2000-04-07
PRIOR APPLICATION NUMBER:	JP 00/280988
PRIOR FILING DATE:	2000-08-03
NUMBER OF SEQ ID NOS:	7059
SOFTWARE:	PatentIn ver. 3.0
SEQ ID NO 4263	
LENGTH:	575
TYPE:	PRY
ORGANISM:	Corynebacterium glutamicum
US - 09738-626-4263	
Alignment Scores:	
Prod. No.:	7.08
Score:	73.00
Percent Similarity:	40.86%
Best Local Similarity:	28.66%
Mismatches:	57
Matches:	47
Conservative:	20
Mismatches:	57

US-09-980-370-5 (1-402) X US-09-738-626-4263 (1-575)

Oy	TTACGCAAGCGTCAGTCGTTTCACATCATGTGACGACACGCGAGATAGCGTACATTCG	67
Db	27 SerAlaValaIaIuSerSerIeuValIleProArglysgInArgIleGlyValPheMet	46
Oy	68 ATGCCGATGGCGATGCGAGCGTTGGCTGCTGTGCTTCGGACA---ATTGAGCTGGC	124
Db	47 AepGIuSerArgIleuSerPheGlyThrAlaGlyLeu-ArgAlaProValGlyProAl	66
Oy	125 TCACAGCTCATTTGGCGGACGACAGTGTAAAGGTC---150	
Db	66 a-----ArgHisIsgImheCasValIeuGlnValThrArgThrThrAlaGlyA	82
Oy	159 -----TTTCTCATGAGATGAGATGGCGTTTGCATTCGATTAACCTGTCTTCGAAG	211
Db	82 IaIaSerTrpLeuAlaGlyIu-ArgAlaAlaIaIeuAsnProValProHisIeuValProGln	102
Oy	212 ATGTTTACAGCGCGACAGCAAGGTAAAGTACATACATCACTTAC-----255	
Db	102 spoliuThGlyIleGlyValGala-----LeuTryProIlnAepGlyProIeuAluArgVal	120
Oy	256 -----GACGGTTTTTCCGGGAGATCAC---PATGGCGTGAATGATTCGCAACAGG	302
Db	120 alValGlyTryAspAlaArgTrycIlySerHisIsthrPheAlaIaIaThrThrAlaGlyValP	140
Oy	302 GCACGACAGCGCGACACATCTACCGTTTTCATCCAGAACAGCCCTAACGGA-----351	
Db	140 heuAlaGlyAlaGly---PheIuValThrIleuLeuProThrProSerProThrProIeuI	150
Oy	352 -----ATAAATTAAGGCGAGACGACAGATTTGGCTGTCTCCACGAGACC-----A	394
Db	159 IeArgTrpIleuValIaSuHyHisGlyIleuAspAlaGlyValGlnIleThrIaIaSerHis	179
Oy	395 ATGCGCGC 402	
Db	179 smuIyAla 181	

```

RESULT 3
US-09-866-050A-303
Sequence: 303, Application US/09866050A
Publication No. US2003004047A1
GENERAL INFORMATION:
Applicant: Watson, James D.
Applicant: Strachan, Lorena
Applicant: Sleeman, Matthew
Applicant: Ornust, Rene
Applicant: Morrison, James G.
Applicant: Kumble, Krishnand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c4
CURRENT APPLICATION NUMBER: US/09/866,050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: PSESD for Windows Version 4.0
SEQ ID NO 303
LENGTH: 617
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-303

Alignment Scores:
Pred. No.:      8.29
Scores:         72.50
Percent Similarity: 41.33%
Best Local Similarity: 29.33%
Query Match:    9.92%
DB:             9
Caps:           3

```

US-09-980-370-5 (1-402) x US-09-866-050A-303 (1-617)

OY 7 TTCACGGAAGCTTCAGTCTTCTACTGAGAGGACGACGAGTGAAGTACTATC 66
DB 301 PheLysPheSerPheHisLysSerHisValGlyProHisLeu 316
OY 67 GATCCGATGCGATGATGAGGATGCTGCTCTCCGACATTTGAGACGCTC 126
DB 317 -----SerGlySerLeuProSerTrpThrValSerMetAlaPheGlnGlnLeu 334
OY 127 ACGCTTATGCGGCGACGACGTAAGCTTTCTCA-----TGCATGGAATGAGCT 180
DB 335 ThrLeuHisPheSerSerTrpGlyThrTyrAlaHisPheSerAlaAlaTrpSerAspProGly 354
OY 181 TTGAATCTGATTAAC-----TATCTCATCTCAAG 210
DB 355 LeuGlyProAspThrProMetProValTyrSerLeuSerGln 369

RESULT 4
US-08-424-550B-82
Sequence 82 Application US/08424550B
Patent No. US2002019447A1
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMM J. PILOT-MATIAS
APPLICANT: GEORGE J. DAMSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: GURUSH R. SCHLAUDER
APPLICANT: ZURUSH R. LENSAT
APPLICANT: ANTHONY SCOTT MUEHROFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIK
APPLICANT: ISA K. MUSHAMMAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
SEQUENCE ID NOS: 1-716
COMPLETION DATE: 1998-08-21
ADDRESSER: ABBOTT LABORATORIES D377/AR6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER: IBM PC COMPATIBLE
MEDIUM TYPE: disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE: 08/08/98
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: POROMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2633
INFORMATION FOR SEQUENCING: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-82

Alignment Scores:
Pred. No.: 13.9 Length: 1422
Score: 71.50 Matches: 29
Percent Similarity: 34.91% Conservative: 8
Best Local Similarity: 27.36% Mismatches: 34

Query Match: 9.78% Indels: 35
DB: 8 Gaps: 7

US-09-980-370-5 (1-402) x US-08-424-550B-82 (1-1422)

OY 30 TATCATGAGAGGAGACGACGACGACGACGACGACGACGACGACGACGACG 89
DB 556 TyrPheTrpAspSerArgThr-----ArgCys---TrpArgTrpSerGly 569
OY 90 TTGCGCTGCTGCTCTCCGACAAATGAGTGGCTTCACAGCTCA---TGGCGGACGAA 146
DB 570 TyrGlyLeuSer-----SerHis***HisPheTrpArgHisLeu 582
OY 147 CGTACGATGCTTCTTCATGATGATGATGATGATGATGATGATGATGATGATG 206
DB 583 CysAlaAlaLeuLeuValTyrTyrIleSerProTyrArg----- 595
OY 207 AAGGATGTTTACAGCGCGACGACGACGACGACGACGACGACGACGACGACG 266
DB 596 -----CysTyrCysArgProSerGly-----***ArgArgGlyAsn 607
OY 267 CGGGATCAGTATGCGGTGATGATCTCAAGACGACGACGACGACGACGACGACG 326
DB 608 ArgGlyGlyValCysIleIleHisSerLeuGlyGly-----HisGly 621
OY 327 TGTTTTCGAAAGAACGCC 344
DB 622 CysGlyAsn***GlnAla 627

RESULT 5
US-09-784-877-2
Sequence 2 Application US/09784877
Patent No. US20010012836A1
GENERAL INFORMATION:
APPLICANT: Exding Hu
APPLICANT: Yuan Zhu
APPLICANT: YUAN ZHU
TITLE OF INVENTION: HUMAN HISTONE DEACETYLASE GENE HD4
TITLE OF INVENTION: GP-70516-C1
CURRENT APPLICATION NUMBER: US/09/784,877
PRIOR APPLICATION NUMBER: 09/138,151
CURRENT FILING DATE: 2001-02-16
PRIOR FILING DATE: 1998-08-21
NUMBER OF SEQ ID NOS: 2
SEQUENCE ID NOS: 1-2
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQUENCE LENGTH: 377
LENGTH: 377
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-784-877-2

Alignment Scores:
Pred. No.: 14.3 Length: 377
Score: 70.00 Matches: 32
Percent Similarity: 33.09% Conservative: 13
Best Local Similarity: 23.53% Mismatches: 45
Query Match: 9.58% Indels: 46
DB: 10 Gaps: 5

US-09-980-370-5 (1-402) x US-09-784-877-2 (1-377)

OY 124 CTACACGCTGATGCGGCGACGACGACGACGACGACGACGACGACGACGACG 183
DB 177 LeuAspLeuHisIleGlyAspGlyValGlnAspAlaPheSerPheThrSerLysValMet 196
OY 184 -----TATCTGATTAATCTCATCTCAAGATGTTTACA--- 219
DB 197 ThrValSerLeuHisLysPheSerProGlyPheProGlyThrIleLysValSerAsp 216
OY 220 ---GGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 267
DB 217 ValGlyLeuGlyLysCysIleArgTyrTyrSerValAsnValProIleGlnAspGlyIleGln 236

Thu Jun 26 11:58:09 2003

us-09-980-370-5.raph

Page 4

[illegible]

```

Db      297 AlaThrLeuIleuGlyGlyGlyTyrAsnLeuAlaAsnTrpAla 312
          |||||               :::: |::::: ||
RESULT 7
US-09-817-538-15
: Sequence 15, Application US/09817538
: Patent No. US2002013712A1
GENERAL INFORMATION:
APPLICANT: J.J. Zuomei
APPLICANT: Bonfills, Claire
INVENTOR: Desiderio Polanco
TITLE OF INVENTION: Desferalase isoforms
FILE REFERENCE: 106101.144
CURRENT APPLICATION NUMBER: US/09/817,538
PRIOR FILING DATE: 2001-03-26
PRIORITY APPLICATION NUMBER: US 60/192,157
PRIORITY FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 3
SEQUENCE NO.: 1
SEQ ID NO. 15
LENGTH: 377
TYPE: PRT
ORGANISM: Human
US-09-817-538-15

Alignment Scores:
Pairwise Score: No.: 14.3 Length: 377
Percent Similarity: 70.00 Matches: 32
Best Local Similarity: 33.09 Conservative: 13
Query Match: 23.53k Mismatches: 45
                9.58% Indels: 46
                Gaps: 5

US-09-980-370-5 (1-402) x US-09-817-538-15 (1-377)
Oy      124 CTCACGCATCTCGGGCCAGCAACCATACGTCTTCATGAGTAAGTACGCTTG 183
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      177 LeuAspLeuHisHisLysLysAspGlyValGlnAspAlaPheSerPheThrSerIleValMet 196
Oy      184 -----AAATCCTGATAACTATCTCATCTCAAGATGTTACA-- 219
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      197 ThrValSerLeuHisLysPheSerProGlyPhePheProGlyThrGlnAspValSerAsp 216
Oy      220 ---GGGCCAACGAACGATCAAGTACATCAC-----TATCAGTCCAAACAGCGCTTTGCC 267
          ||| ||| ||| :||||| ||| ||| ||| ||| ||| ||| ||| |||
Db      217 ValGlyLeuGlyLysLysLysGlyTyrSerValAsnValProIleGlnAspGlyIleGln 239
Oy      268 GGAGATCACTAT-----||| 276
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      237 AspGluValGlyTyrGlnIleCysGlnSerValLeuLysGluValTyrGlnAlaPheAsn 256
Oy      280 -----GGCGATGATCATCTCCAAGACGGCCAGCAACCGCGAGCATTCACGGTTGTT 330
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      257 ProLysAlaValAlaValLeuGlnGlnLeuGlnValAspThrIleAlaGlyAspPromecysSer 276
Oy      331 TTCGAGAAGAACGCGCTTACGGAATTAATTAAG----- 360
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      277 PheAsnMetThrProValGlyIleGlyLysCysLeuLysTyrIleLeuGlnIntArgInLeu 296
Oy      361 -----GGCGAGCAAGATTCCGCTCTTCCACGCAAGGCC 393
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      297 AlaThrLeuIleuGlyGlyGlyTyrAsnLeuAlaAsnTrpAla 312

RESULT 8
US-09-308-207-33
: Sequence 33, Application US/09308207
: Patent No. US2003022323A1
GENERAL INFORMATION:
APPLICANT: MARIA DIAZ-TORRES ET AL.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT PRODUCTION OF 1,3 PROBANDIOL
NUMBER OF SEQUENCES: 68

```



```

OY 136 GCGCGCACCAACAGTATAGATCTTTTCATGAGATGGAATGGCTTTGATCTGATTAAC 195
    ||| : : :
DB 636 GlysanserGlnIle----- 640
OY 196 TATCTCATCTCAAGAGATGTTACAGCGGCAAGGTAAGTACTGATCTGATTAAC 255
    ||| : : :
DB 641 ----- 648
OY 256 GACGCTTTCCCGGAGTACTATGCGGTGATGATCTCAAGACGCGGACGACGCGCGA 315
    : : : ||| : : :
DB 649 GlnMetPhe-----GlyTyrAlaIthrValLeuArgSerAlaSerGlnGlyArgGly 665
    ||| : : : ||| : : :
OY 316 GACTTCACAGCTTGTTC-----GAGCAACGCGCTTAAGCAATA----- 354
    ||| : : : ||| : : :
DB 666 ThrPheMetMetValPheAspHisTyrGlnAspValProLysSerValGlnGlnGlnIle 685
OY 355 -----AATTAAGGCG 363
    ||| : : :
DB 686 IleLysLysAsnLysGly 691

RESULT 12
US-10-125-692-18
: Sequence 18, Application US/10125692
: Publication No. US2003004429A1
: GENERAL INFORMATION:
: APPLICANT: Adarem, Alan
: APPLICANT: Hayashi, Fumitaka
: APPLICANT: Smith, Kelly D.
: APPLICANT: Underhill, David M.
: APPLICANT: Ozlinsky, Adrian
: TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
: FILE REFERENCE: P-15 5155
: CURRENT APPLICATION NUMBER: US/10/125,692
: PRIOR FILING DATE: 2002-04-17
: PRIOR APPLICATION NUMBER: US 60/285,477
: PRIOR FILING DATE: 2001-04-20
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 18
: LENGTH: 506
: TYPE: PRF
: ORGANISM: S. typhimurium2
US-10-125-692-18

Alignment Scores:
Pred. No.: 31 Length: 506
Score: 67.50 Matches: 36
Percent Similarity: 37.88% Conservative: 14
Best Local Similarity: 27.27% Mismatches: 53
Query Match: 9.23% Indels: 29
DB: 9 Gaps: 6

US-09-980-370-5 (1-402) x US-10-125-692-18 (1-506)
OY 58 ACTACTATGATGCCGATGCGGATGCGAGGCTGCTGCTCTCGGACATTTG 117
    ||| : : :
DB 196 ThrThrIleAspValSerGlyLeuAspAlaIle----- 208
OY 118 GACTGGCTCACAGCTCATGCGGCGACCAAGCTAGTACTGTTCTCATGGAATGATG 177
    ||| : : : ||| : : :
DB 209 -----LysAlaIleAlaIthrGlyGlyThrAsnGlyThrAlaSerValIThrGlyAlaVal 226
    ||| : : : ||| : : :
OY 178 GCTTTGATCTCGATTAAC-----TATCTCATCTCAAGAGTGTTCACAGCGCGACG 228
    : : : ||| : : :
DB 227 LysPheAspAlaAspAsnAsnLysTyrPheValIThrIleGlyGlyPheThrGlyAlaAsp 246
    ||| : : : ||| : : :
OY 229 AAGGTAAGTAC---TACTATCCATCAACGACGCGGTTTCCCGGGGAT---CACTATGCG 282
    ||| : : : ||| : : :
DB 247 AlaAlaLysAsnGlyAspTyrGlnValAlaSerValAlaIthrAspGlyThrValIthrLeuAla 266
    ||| : : : ||| : : :
OY 283 GTCATGATCTCCAAAGCGGCGACGAGACCGCGGAGACTTCACG----- 324

```

```

DB 267 AlaGlyAlaIthrLysThrThrMetProAlaGlyAlaIthrLysThrGlnValGlnGlu 286
    : : : ||| : : :
OY 325 -----GTTGTTTGGAAAGAACGCTTACGAGTAATAATGAAGCGCGA 366
    ||| : : : ||| : : :
DB 287 LeuLysAspThrProAlaValAlaSerAlaAspAlaLysAsnAlaValIleAlaGlyGly 306
    ||| : : : ||| : : :
OY 367 GCAGATTTCGCTTTCACGGAAGCCCAATGGCGCC 402
    : : : ||| : : :
DB 307 ValAsp-----AlaThrAspAlaAsnGlyAla 315

RESULT 13
US-09-980-748-1745
: Sequence 1745, Application US/09880748
: Publication No. US20030059937A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: Antipodites that Immunospecifically Bind BlyS
: FILE REFERENCE: PE523
: CURRENT APPLICATION NUMBER: US/09/880,748
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: 60/212,210
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: 60/240,816
: PRIOR FILING DATE: 2000-10-17,248
: PRIOR APPLICATION NUMBER: 60/276,248
: PRIOR FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/277,379
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/293,499
: PRIOR FILING DATE: 2001-05-25
: NUMBER OF SEQ ID NOS: 3239
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1745
: LENGTH: 256
: TYPE: PRF
: ORGANISM: Homo sapiens
US-09-980-748-1745

Alignment Scores:
Pred. No.: 29.2 Length: 256
Score: 67.00 Matches: 31
Percent Similarity: 33.12% Conservative: 21
Best Local Similarity: 19.75% Mismatches: 49
Query Match: 9.17% Indels: 56
DB: 9 Gaps: 7

US-09-980-370-5 (1-402) x US-09-980-748-1745 (1-256)
OY 46 CCAGCGAATGAGTACTATGATGCGGATGCGAGGTTGCTGCTGCTGCT 105
    ||| : : : ||| : : :
DB 14 ProSerGlnThrIleSerLeuThrCysThrValSerGlyGlySer-----IleSer 30
OY 106 TCCGACCAATG-----GACTGGCTCACAGCTCATGAGCGGCGACCAAGTAACTCT 159
    ||| : : : ||| : : :
DB 31 SerGlyAsnTyrTyrTTPSerTTPValArgGlnIleSerProly-----LysGly 46
    ||| : : : ||| : : :
OY 160 TTCTCATGGAATGCA----- 174
    ||| : : :
DB 47 LeuGlnTTPIleGlyTyrIleTyrAspIleGlyAsnThrTyrAsnProSerLeuLysSer 66
    ||| : : : ||| : : :
OY 175 -----ATGGCTTGAATCTGATTAACATATCTATCTCAAGGATGTTCACAGCGCGA 225
    ||| : : : ||| : : :
DB 67 ArgValIthrMetSerValAspThrSerLysAsnGlnIleSerLeuGlnLeuThrSerVal 86
    ||| : : : ||| : : :
OY 226 ACG-----AAGTAAAGTACTACTATCCAGTCAC 255
    ||| : : : ||| : : :
DB 87 ThrAlaAlaAspThrAlaValTyrTyrCysAlaArgValProTyrTyrTyrAspThrSer 106
    ||| : : : ||| : : :
OY 256 GACGCTTTCCCGGAGTACTAT-----GCGGTG 285
    ||| : : : ||| : : :
DB 107 GlyGlyTyrLeuGlnGlyGlnTyrTyrTyrGlyMetAspValIleProGlyGlnGlyThrLeuVal 126

```

QY 286 ATGATCTCCAGACGGCGACGACGAGACTTACAGGTGTGTTTCGAGAAACGCT 345
: : : : :
Db 127 ThrValSerSerGlyGlyGlySerGlySerGly----- 137
QY 346 AAGCGATTAATTAAGGCGAGACGATTCGCTTTCACGAGCAAT 396
: : : : :
Db 138 GlyGlySerGlyGlyGlySerGlySerGlyValLeuThrGlnProAlaSer 154
RESULT 14
US-09-371-900-40
: Sequence 40, Application US/09371900
: Patent No. US20020137700A1
: GENERAL INFORMATION:
: APPLICANT: FALB, DEAN A
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/371,900
: FILING DATE: 11-Aug-1999
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/599,654
: FILING DATE: 09-FEB-1996
: APPLICATION NUMBER: US 08/485,573
: FILING DATE: 07-JUN-1995
: APPLICATION NUMBER: US 08/386,844
: FILING DATE: 10-FEB-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: CORUZZI, LAURA A
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7853-104
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 40:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1481 amino acids
: TYPE: amino acid
: STRANDEDNESS: <unknown>
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-371-900-40
Alignment Scores:
Pred. No.: 48.5 Length: 1481
Score: 67.00 Matches: 26
Percent Similarity: 38.46% Conservatve: 9
Best Local Similarity: 28.57% Mismatches: 28
Query Match: 9.54% Indels: 28
DB: 10 Gaps: 3
US-09-980-370-5 (1-402) x US-09-371-900-40 (1-1481)
QY 233 ACCCTGCTGGCGCGCTTAACATCTTGAGATGAGATAGTATCAGAGTTCGAAGCAAT 174
: : : : :
Db 1012 ThrGlyLeuIleIleProIleuThrSer-----Val 1020
QY 173 CCATTCCATGAGAAAGACTTACTAGTTGTGTCGCCCATGAGCTGTGAGCCAGTCCAAAT 114

Db 1021 ProThrSerAlaGlySerGluMetThrThrIleu-----GlyValThrAlaGluIleu 1037
QY 113 TGTCCGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 75
: : : : :
Db 1038 SerProIleSerIleSerIleuGlyThrSerProSerProGlnThrThrValIleSerThr 1057
QY 74 -----TCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 36
: : : : :
Db 1058 AlaGluAspLeuAlaProIleSerAlaThrIleAlaValGlnSerSerThrGlnSerPro 1077
QY 35 TGAGTAGAAGACTCGACGATTCGCGTCCGAGAGCT 3
: : : : :
Db 1078 ThrThrLeuSerSerSerAlaSerValAlaSer 1088
RESULT 15
US-09-924-417-60
: Sequence 60, Application US/09924417
: Patent No. US2002014241A1
: GENERAL INFORMATION:
: APPLICANT: Falb, Dean
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR
: DISEASE
: NUMBER OF SEQUENCES: 67
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/924,417
: FILING DATE: 07-Aug-2001
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/09/034,286
: FILING DATE: 04-MAR-1998
: APPLICATION NUMBER: 08/870,434
: FILING DATE: 06-JUN-1997
: APPLICATION NUMBER: 08/799,910
: FILING DATE: 13-FEB-1997
: APPLICATION NUMBER: 60/011,787
: FILING DATE: 16-FEB-1996
: APPLICATION NUMBER: 08/599,654
: FILING DATE: 09-FEB-1996
: APPLICATION NUMBER: 08/485,573
: FILING DATE: 07-JUN-1995
: APPLICATION NUMBER: 08/386,844
: FILING DATE: 10-FEB-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: CORUZZI, LAURA A
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7853-114-999
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)7909090
: TELEFAX: (212)8699741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1481 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-924-417-60

Alignment Scores:

Pred. No.:	48.5	Length:	1481
Score:	67.00	Matches:	26
Percent Similarity:	38.46%	Conservative:	9
Best Local Similarity:	28.57%	Mismatches:	28
Query Match:	9.54%	Indels:	28
DB:	10	Gaps:	3

US-09-980-370-5 (1-402) x US-09-924-417-60 (1-1481)

```
QY 233 ACCTTCGTCGCGCCTGTACATCCTTTGAGATGAGATAGTATCATCAAGCCATT 174
    ||| ::::|:|:|:|:|:|
Db 1012 ThrglyleuileProleuthrSer-----Val 1020

QY 173 CCATTCCATGAGAAAGACCTTACTACGTTGGTCCGCCCATGAGCTGTGAGCCACGCCAAT 114
    ||| |:|:|:|:|:|:|:|
Db 1021 ProthrSerAlaLysGluMetThrThrLysLeu-----GlyValThrAlaGluTyr 1037

QY 113 TGTCCGAGAGACAGACAGACCACCTTCACCATCGCCA----- 75
    ||| |:|:|:|:|:|
Db 1038 SerProAlaSerArgSerLeuGlyThrSerProSerProGluThrThrValValSerThr 1057

QY 74 -----TCGSCATGATAGTACATTCGGCTGCTGCTGCTGCCA 36
    ||| |:|:|:|:|:|:|
Db 1058 AlaGluAspLeuAlaProLysSerAlaThrPheAlaValGlnSerSerThrGlnSerPro 1077

QY 35 TGAGTAGAAGACTCGAACCTTCCGCGAAGTCT 3
    |||:|:|:|:|:|
Db 1078 ThrThrLeuSerSerSerSerAlaSerValAlaSer 1088
```

Search completed: June 26, 2003, 00:50:06
Job time : 75.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 26, 2003, 00:32:09; Search time 37 seconds

(without alignments)
2088,975 Million cell updates/sec

Title: US-09-980-370-5

Sequence: 1 gcagacttcacgaaacgtl.....ccaggaagcgaatgcgc 402

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_n2p.model -DEV=11h
-O=/cgn2_1/USPRO.spool/US09980370/runat_19062003-174443-8865/app.query.fasta.1.583
-R=3 -O=frame_plus_n2p.model -X=1000000000 -X=1000000000
-UR=1000000000 -UR=1000000000 -UR=1000000000
-DOCALIGN=200 -THR.SCORE=100 -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=100 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09980370.ecgn.1.1.25.8runat_19062003-174443-8865 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQDERR -NEG_SCORES=0 -WATR -DSPLUCK=100 -LONGLOG
-DEXT=100 -WATR -DSPLUCK=100 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7
Database: PIR-73:*

1: PIR-73:
2: PIR-73:
3: PIR-73:
4: PIR-73:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	716	97.9	1704	A55426	gingipain R (EC 3.1.3.1)
2	715	97.8	2632	A55426	lysine-specific cy
3	715	97.8	2632	A55426	gingipain R (EC 3.1.3.1)
4	706	96.6	1526	B81094	isomerase, probab
5	90.5	12.4	218	B81094	probable 2-oxo-1-
6	90.5	12.4	218	B81094	hypothetical prote
7	89	12.2	418	D70154	probable hydrolase
8	88.5	12.1	218	C81947	cellulase (EC 3.2.1.4)
9	83.5	11.6	429	A55426	hypothetical prote
10	83.5	11.6	429	A55426	hypothetical prote
11	82.5	11.4	373	B70819	probable carotene
12	82.5	11.4	373	B70819	hypothetical prote
13	81.5	11.1	966	C69189	hypothetical prote
14	80	10.9	759	B84481	cycloartenol synth

15	79.5	10.9	548	1	HYB57
16	78.5	10.9	551	2	B3706
17	78.5	10.9	551	2	B3706
18	78.5	10.9	551	2	B3706
19	78.5	10.9	551	2	B3706
20	77.5	11.0	3014	1	JC5620
21	77.5	11.0	1102	2	A3247
22	76	10.3	527	2	PN0114
23	75.5	10.3	187	2	E61834
24	75.5	10.3	479	2	A69832
25	75.5	10.3	607	2	S49528
26	75.5	10.3	1034	2	A36108
27	75.5	10.3	759	2	A49398
28	75	10.3	985	2	T29910
29	74.5	10.2	556	2	T29910
30	74.5	10.2	556	2	T29910
31	74.5	10.2	556	2	T29910
32	74	10.1	453	2	B49022
33	74	10.1	504	2	S54744
34	74	10.1	554	2	G72361
35	74	10.1	666	2	E71310
36	74	10.1	2114	2	T29910
37	73.5	10.1	544	2	A49464
38	73.5	10.1	546	2	JC4113
39	73.5	10.1	546	2	S72176
40	73.5	10.1	546	2	S72176
41	73.5	10.1	640	2	S72175
42	73.5	10.1	640	2	B81423
43	73.5	10.1	435	2	S72175
44	73	10.1	435	2	S72175
45	73	10.0	435	2	JC7137

ALIGNMENTS

Result 1
A55426
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP
C:Species: Porphyromonas gingivalis
C:Date: 10-Feb-1995 (sequence, revision 10-Feb-1995) (text, change 08-Oct-1999)
C:Accession: A55426, D55113, P196, R.N.: Prochazka, V., Kiefer, M.C., Travis, J., Bart
J. Biol. Chem. 270, 1007-1010, 1995
A:Title: Molecular cloning and structural characterization of the Arg-gingipain prote
A:Reference number: A55426; MUID:95138080; PMID:7836351
A:Accession: A55426
A:Status: preliminary
A:Keywords: preliminary
A:Cross-references: GB:U15282; NID:9557067; PID:AAA69539.1; PID:9557068
J. Biol. Chem. 269, 406-411, 1994
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso
A:Accession: A55113; MUID:9410243; PMID:8278827
A:Status: preliminary
A:Keywords: preliminary
A:Molecule type: protein
A:Residues: 228-249 (PIR)
A:Note: sequence extracted from NCBI backbone (NCBIP:141694)
A:Keywords: cysteine proteinase; hydrolase
Alignment Scores:
Pred. No.: 3,956-68
Score: 716.00
Percent Similarity: 99.28
Best Local Similarity: 97.98
Identical: 2
Gaps: 0
US-09-980-370-5 (1-402) x A55426 (1-1704)

OY 1 GCAGACTTCACGGAACGGTTCGAGTCTTCACTGATGAGAGGACGACGCGAATGGACT 60
 |||||
 Db 1137 AAlasphethergulturphecgluuserterthrhslgylualaproalaglutprphr 1156
 OY 61 ACTATGATGCCGATGCCGATGCGAGAGGCTGGCTCTGCTCTTCGCGAATGGAC 120
 |||||
 Db 1157 Thilleasphalaspclaspclaspclaspclaspclaspclaspclaspclaspclasp 1176
 OY 121 TGGCTACAGCTGATGCGGACACAGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 180
 |||||
 Db 1177 TTPleutrhralahslgylgylthrasnvalvalalaserphesertrpasmnglymelala 1196
 OY 181 TTGATCTGCTGATCTCATCTCAAGAGTGTTCACGCGGACGCGAAGGTAAGTAC 240
 |||||
 Db 1197 Leuaspromaspasntylleuilerlvaspvalthrglyalathrlvsvallystyr 1216
 OY 241 TACTATCCAGTCAACGAGCTTTTCGCGGATCAGTATGCGATGATGATGATGATGATGATGAT 300
 |||||
 Db 1217 Tyrtlyalavalaasphclaspclaspclaspclaspclaspclaspclaspclaspclasp 1236
 OY 301 GCGCAGAACGCGGACGACTTCACGAGTGTTCGCGAAGGCGTACGCGAATGAATAG 360
 |||||
 Db 1237 Glythrasnvalaspclaspclaspclaspclaspclaspclaspclaspclaspclasp 1256
 OY 361 GCGGAGCAAGATTCGCTTCGCGGACGCGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 402
 |||||
 Db 1257 Glylyalavalaasphclaspclaspclaspclaspclaspclaspclaspclaspclasp 1270

RESULT 2

T30836
 Lysine-specific cysteine proteinase porphyryn (EC 3.4.22.-) - Porphyromonas gingivalis
 N:Alternate names: Lysine-specific cysteine proteinase 1, 60K
 C:Species: Porphyromonas gingivalis
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 A:Accession: T30836; T30837; T30526; A53113
 J:Biochem. 179: 2734-2741, 1996
 A:Title: Analysis of the prt gene encoding porphyryn, a cysteine proteinase of Porphy-
 A:Reference number: 220895; PMID:86213011; PMID:8631659
 A:Accession: T30836
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-795
 A:Cross-references: EMBL:U42210; NID:9134325; PID:9134326; PIDN:AA06565.1
 R:Slakecki, N.; Cleel, S.M.; Reynolds, E.C.
 A:Submitted to the EMBL Data Library, October 1996
 A:Reference number: 220896
 A:Accession: T30837
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-795
 A:Cross-references: EMBL:U75366; NID:92182811; PID:92182812; PIDN:AA060809.1
 A:Title: The hemagglutinin gene A (hga) of Porphyromonas gingivalis 381 contains fou
 A:Accession: T28651
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2628
 A:Cross-references: EMBL:U41807; NID:91552410; PID:91469916; PIDN:AA017128.1
 C:Genetics:
 A:Gene: hga
 Alignment Scores:
 Pred. No.: 1,426-67 Length: 2628
 Score: 711.00 Matches: 130
 Percent Similarity: 99.254
 Query Match: 97.254
 DB: 2 Gaps: 0
 US-09-980-370-5 (1-402) x T28651 (1-2628)
 OY 1 GCAGACTTCACGGAACGGTTCGAGTCTTCACTGATGAGAGGACGACGCGAATGGACT 60
 |||||
 Db 685 AAlasphethergulturphecgluuserterthrhslgylualaproalaglutprphr 704
 OY 61 ACTATGATGCCGATGCCGATGCGAGAGGCTGGCTCTGCTCTTCGCGAATGGAC 120

C:Keywords: cysteine proteinase; hydrolase

Alignment Scores:

Pred. No.: 5,066-68 Length: 1732
 Score: 715.00 Matches: 131
 Percent Similarity: 99.254
 Query Match: 97.814
 DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x T30836 (1-1732)

OY 1 GCAGACTTCACGGAACGGTTCGAGTCTTCACTGATGAGAGGACGACGCGAATGGACT 60
 |||||
 Db 1157 AAlasphethergulturphecgluuserterthrhslgylualaproalaglutprphr 1176
 OY 61 ACTATGATGCCGATGCCGATGCGAGAGGCTGGCTCTGCTCTTCGCGAATGGAC 120
 |||||
 Db 1177 Thilleasphalaspclaspclaspclaspclaspclaspclaspclaspclaspclasp 1196
 OY 121 TGGCTACAGCTGATGCGGACACAGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 180
 |||||
 Db 1197 TTPleutrhralahslgylgylthrasnvalvalalaserphesertrpasmnglymelala 1216
 OY 181 TTGATCTGCTGATCTCATCTCAAGAGTGTTCACGCGGACGCGAAGGTAAGTAC 240
 |||||
 Db 1217 Tyrtlyalavalaasphclaspclaspclaspclaspclaspclaspclaspclaspclasp 1236
 OY 241 TACTATCCAGTCAACGAGCTTTTCGCGGATCAGTATGCGATGATGATGATGATGATGATGAT 300
 |||||
 Db 1237 Glythrasnvalaspclaspclaspclaspclaspclaspclaspclaspclaspclasp 1256
 OY 301 GCGCAGAACGCGGACGACTTCACGAGTGTTCGCGAAGGCGTACGCGAATGAATAG 360
 |||||
 Db 1257 Glythrasnvalaspclaspclaspclaspclaspclaspclaspclaspclaspclasp 1276
 OY 361 GCGGAGCAAGATTCGCTTCGCGGACGCGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 402
 |||||
 Db 1277 Glylyalavalaasphclaspclaspclaspclaspclaspclaspclaspclaspclasp 1290

RESULT 3

T28651
 Hemagglutinin A - Porphyromonas gingivalis
 N:Alternate names: Hemagglutinin A
 C:Species: Porphyromonas gingivalis
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Aug-2001
 A:Accession: T28651
 R:Han, N.; Whitlock, J.; Proguiske-Fox, A.
 A:Infect. Immun. 64: 4000-4007, 1996
 A:Title: The hemagglutinin gene A (hga) of Porphyromonas gingivalis 381 contains fou
 A:Accession: T28651
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2628
 A:Cross-references: EMBL:U41807; NID:91552410; PID:91469916; PIDN:AA017128.1
 C:Genetics:
 A:Gene: hga
 Alignment Scores:
 Pred. No.: 1,426-67 Length: 2628
 Score: 711.00 Matches: 130
 Percent Similarity: 99.254
 Query Match: 97.254
 DB: 2 Gaps: 0
 US-09-980-370-5 (1-402) x T28651 (1-2628)
 OY 1 GCAGACTTCACGGAACGGTTCGAGTCTTCACTGATGAGAGGACGACGCGAATGGACT 60
 |||||
 Db 685 AAlasphethergulturphecgluuserterthrhslgylualaproalaglutprphr 704
 OY 61 ACTATGATGCCGATGCCGATGCGAGAGGCTGGCTCTGCTCTTCGCGAATGGAC 120

Db 705 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerLeuSerGlyGlnLeuAsp 724
 QY 121 TGGCTACAGCTCAGTGGCGGACCCAGCAGTGTAGCTCTTTCATGCAATGGAATGGCT 180
 Db 725 TrpLeuThrAlaIleAspGlyGlyThrAsnValValAlaSerPheSerThrAsnGlyMetAla 744
 QY 181 TTGAATCCTGTACTATCTCATCTCAACAGATGTACAGCGGCAACGAAAGTAAGTAC 240
 Db 745 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 764
 QY 241 TACTATCCAGTCAACGACGGCTTTCCGGGATCACTATGCGGTGATGATCTCAAGACG 300
 Db 765 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 784
 QY 301 GCCACGAGACCGCGGAGCTTCACGGTGTTCGAAAGAAAGCGCTTAAGGATTAATAG 360
 Db 785 GlyThrAsnAlaGlyAspPheThrValValPheGlnGlnThrProAsnGlyIleAsnLys 804
 QY 361 GCCGAGACGAGATTCGCTCTTCCAGCGAAGCCAAATGGCGCC 402
 Db 805 GlyGlyAlaAspArgPheGlyLeuSerThrGluAlaAspGlyAla 818

RESULT 4

glingidain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)
 C:Species: Porphyromonas gingivalis
 C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997
 R:Accession: S49763
 R:Aduse-Opoku, J.; Muir, J.; Stanley, J.M.; Rangarajan, M.; Curtis, M.A.
 submitted to the EMBL Data Library, November 1994
 A:Description: Cloning, sequence analysis and expression in *Escherichia coli* of ppr1 of
 A:Reference number: S49763
 A:Accession: S49763
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1526 <ADU>
 A:Cross-references: EMBL: X82680
 C:Genetics:
 A:Gene: ppr1
 C:Keywords: cysteine proteinase; hydrolase

Alignment Scores:
 Pred. No.: 4,58e-67 Length: 1526
 Score: 706.00 Matches: 130
 Percent Similarity: 98.51% Conservative: 2
 Best Local Similarity: 97.01% Mismatches: 0
 Query Match: 96.58% Indels: 0
 Gaps: 0

US-09-980-370-5 (1-402) x S49763 (1-1526)

QY 1 GGAGACTTCACGGAAAGCTTCGACTCTTACTCATGAGAGGACACCGAATGACT 60
 Db 1139 AlaAspPheThrGlnThrPheGluSerSerThrHisGlyGluAlaThrLysValLysThr 1158
 QY 61 ACTATGATGCGGATGGGATGTGAGGCTGTCTGTCTCTCCGCAATGTGAC 120
 Db 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerLeuSerGlyIleLeuAsp 1178
 QY 121 TGGCTACAGCTCAGTGGCGGACCCAGCAGTGTAGCTCTTTCATGCAATGGAATGGCT 180
 Db 1179 TrpLeuThrAlaIleAspGlyGlyThrAsnValValAlaSerPheSerThrAsnGlyMetAla 1198
 QY 181 TTGAATCCTGTACTATCTCATCTCAACAGATGTACAGCGGCAACGAAAGTAAGTAC 240
 Db 1199 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1218
 QY 241 TACTATCCAGTCAACGACGGCTTTCCGGGATCACTATGCGGTGATGATCTCAAGACG 300
 Db 1219 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1238
 QY 301 GCCACGAGACCGCGGAGCTTCACGGTGTTCGAAAGAAAGCGCTTAAGGATTAATAG 360

Db 1239 GlyThrAsnAlaGlyAspPheThrValValPheGlnGlnThrProAsnGlyIleAsnLys 1258
 QY 361 GCCGAGACGAGATTCGCTCTTCCAGCGAAGCCAAATGGCGCC 402
 Db 1259 GlyGlyAlaAspArgPheGlyLeuSerThrGluAlaAspGlyAla 1272

RESULT 5

Isomerase, probable NMB1338 [imported] - *Neisseria meningitidis* (strain MC58 serogroup B)
 C:Species: *Neisseria meningitidis*
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 R:Accession: B81094
 R:Rettlino, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; Science 287, 1809-1815, 2000
 A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A:Reference number: AB1000; MUID:20175755; PMID:10710307
 A:Accession: B81094
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <TEP>
 A:Cross-references: GB:AE002482; GB:AE002098; NID:g7226577; PIDN:AAF41713.1; PID:g722
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1338
 C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Alignment Scores:
 Pred. No.: 0.225 Length: 218
 Score: 90.50 Matches: 38
 Percent Similarity: 36.36% Conservative: 14
 Best Local Similarity: 26.57% Mismatches: 48
 Query Match: 12.36% Indels: 43
 Gaps: 6

US-09-980-370-5 (1-402) x B81094 (1-218)

QY 67 GATGCCGATGGCGGATCGAGGCT----- 90
 Db 77 AspSerAspGlyThrGlnGlyGlnGlyLysAspIleLeuGlnCysValAlaIleGlyTrpGlyVal 96
 QY 91 -----TGGCTGTCTGTCTCTCCGCAATGTGACTGG 123
 Db 97 GlyLeuAspPheThrAlaAspArgIleGlnCysArgLeuLysGlnLysGlyLeuProTrp 116
 QY 124 CTCACAGCTCATGGC-----GCCACCAACGTAGTAAGCTTTTCATGCAATGGAATG 177
 Db 117 LeuLysAlaIleLysGlyPheArgHisSerAlaCysValSerAspPheAlaIleGlyArg 136
 QY 178 GCTTGATTCGATACATATCATCTCAACAGATGTTACAGCGGCAACGAAAGTAAG 237
 Db 137 IleGlyAsnProGlnGlyValLeuPheSerLeuLysGlnAsnGlyValLeuLysGlnArg 156
 QY 238 -----TACTACTATCCAGTCAACGAGCGGTTTCCCGGATCATATGCG 282
 Db 157 GlyAspThrGlyLeuMetIleThrProIleArgGlu----- 168
 QY 283 GTGATGATCTCCAGACG-----GCCACCAACCGCGGAGACTTCACGTTGTT 330
 Db 169 ----IleLeuHisLysLeuAlaAlaAspTyrGlyLeuGlnLysGlyAsp-----LeuVal 185
 QY 331 TTGGAAGAAAGCGCTTAAGGATTAATAGCGGAGCAAGATTCGTTCTTCACAGGAA 390
 Db 186 PheThrGlyThrProSerGlnValGlyAlaIleGlyAlaGlyAspAsnLeuAlaIleGln 205
 QY 391 GCCAATGGC 399
 Db 206 LeuAspGly 208

RESULT 6

QY	124	CTCACAGCTCAATGGC-----GGCACCAACAGTAGTAAAGCTCTTCATCATGATGGCAATG	177
Db	117	LeuLysAlaLeuLysGlyPheArgHisSerIaAcysValSerAspPheIaIaIaIaGlyArg	166
QY	178	GCTTTGAATCCTCGATAACTATCTCATCTCAAGAGATGTTACAGCGCCACCAAGAGTAAG	237
Db	137	IleGlyAsnProGlyAsnIValLeuPheSerLeuLysGlnAsnGlyAlaLeuLysGlnArg	156
QY	238	-----TACTACTATCCACGACCAACGAGGTTTTCCCGGGATCACTATCGC	282
Db	157	GlyAspThrGlyLeuMetIleTyrProIleArgGlu-----	168
QY	283	GTGATGATCTCCACAGC-----GGCACCAACGCCGAGACTTCACGGTTGTT	330
Db	169	--IleLeuHisLysLeuAlaIaIaAspTyrGlyLeuGlyLysGlyAsp-----LeuVal	185
QY	331	TTTGCAGAAACCCCTTAACGGAATAAATAAGGCGCGAGCAAGATTCCGCTTTTCCAGCGAA	390
Db	186	PheThrGlyThrHisSerSerGlyValGlyAlaIleGlyIaIaGlyAspAsnLeuAlaLeuGlu	205
QY	391	GCCCAATGCC	399
Db	206	LeuAspGly	208

QY 334 GAAGAACGCTAACGGAAATAAATAGCGCGAGCAAGATTGCGTCTTCCACGGAGCC 393
Db 474 AsnLysAlaAlaTyrLeuIleSerGlnGlyThrIstYrGlyAlaSerValGly 493
QY 394 AATGGC 399
Db 494 IleGly 495

RESULT 11
B70819
probable carnitine operon oxidoreductase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70819
R:COLE, S.T.; BROSCH, R.; PARKHILL, J.; GARNIER, T.; CHURCHER, C.; HARRIS, D.; GORDON, R.; COMOR, R.; DAVIES, R.; DEVLIN, K.; FELLWELL, T.; GENTLES, S.; HAMLIN, N.; HOLROYD, S.; REJANDREEM, M.A.; ROGERS, J.; RUTER, S.; SEEGER, K.; SKELTON, S.; SQUARES, S.
NATURE 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70819
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-373 <COL>
A:Cross-references: GB:AL022002; GB:AL123456; NID:93261544; PIDN:CAAI7610.1; PID:9326154
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fade16
C:Superfamily: acyl-CoA dehydrogenase

Alignment Scores:
Pred. No.: 1.5E Length: 373
Score: 83.00 Matches: 39
Percent Similarity: 39.13% Conservative: 15
Best Local Similarity: 28.26% Mismatches: 64
Query Match: 11.35% Indels: 20
Gaps: 2 6

US-09-980-370-5 (1-402) x B70819 (1-373)

QY 25 TCTTCTACTCATGAGACGACCGAGGATGACTATCGATGCGGATGCGATGCT 84
Db 118 SerArgSerHisPheThrPheAlaProVal-----SerThrAlaSerAlaAspGly 135
QY 85 GAGGGTGGCTGCTGCTGCTTCCGACAAATTGAGTGGCTCACAGCTCATGGCGGAC 144
Db 136 -----IleAlaValAlaGAlaAspLysSerTrpValThrSerAlaGlyPheAla 151
QY 145 AAC-----GTAGTAGCTCTTTCATGATGAGATGGCTTGATCCGATAACTAT 198
Db 152 AspValTyrValValSerValGlySerAlaAspGlyAlaIleGlyAspValAspLeuTyr 171
QY 199 CTCATCTCAAGAGGATTTACAGCGCGCAACGAGGTAAGTACTATTCAGTCAAGCAGC 258
Db 172 AlaValAlaProAlaAspThrProGlyLeuArgValAlaGlyThrPhe-----ThGlyMet 189
QY 259 GGTTTTCCGGGAGTACTATGCGGTGATGATC-----TCC 294
Db 190 GlyLeuArgGlyAsnAlaSerAlaProMetAlaValAlaSerIleArgIleProAspSerTyr 209
QY 295 AAGAGCGGACGACGAAACCGGAGACTTCAGGTTGTTTCGAGAAAGCGCTAAGGAAATA 354
Db 210 ArgLeuGlyGlyAlaIleGlyGlyGlyPheGlyIleMetLeuGlnThrValLeuProTrpPhe 229
QY 355 AATAAGGCGGAGACA-----AGATTGCGTCTTTCACGAGCAAGCAATGGCGCC 402
Db 230 AsnLeuGlyAsnAlaIleValSerLeuGlyLeuAlaIleThrAlaIleGlyAla 247

RESULT 12
A99283
hypothetical protein dppc-2 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: A99283
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
submitted to GenBank, May 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A99283
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <XUN>
A:Cross-references: GB:AE006641; NID:913814480; PIDN:AAK41520.1; GSPDB:GN00155
C:Genetics:
A:Gene: dppc-2

Alignment Scores:
Pred. No.: 1.78 Length: 536
Score: 82.50 Matches: 41
Percent Similarity: 38.89% Conservative: 22
Best Local Similarity: 25.31% Mismatches: 34
Query Match: 11.29% Indels: 65
Gaps: 2 10

US-09-980-370-5 (1-402) x A99283 (1-536)

QY 76 GCGATGGTACGCGTTCGCTGCTGCTTCCGACAAATTGGACTGCTCACAGCTCAT 135
Db 164 GlyAspTyrGluPheTyrIleIleGlnAlaSer-----IleSerAlaTyr 178
QY 136 GCGGCGACCAAGCTGTAGTACCTCTTCATGAGATGAATGAGCTTGAATCTGAT--- 192
Db 179 ---AlaThrAsn-----AlaTyrTyrThrPheAsnGlyTyrValIleAsnProSerAsn 195
QY 193 -----AACTATCTCATCTCAAGAGATGTACAGCGCA 225
Db 196 AlaThrPheTrpLeuPheLeuAlaGlyAsnTyrLeuProThrAsnLeuValThrLeuSer 215
QY 226 ACGAAGCTAAATAC----- 240
Db 216 ThrValAlaPheLysTyrLeuGlyAsnGlyIlyTrpAsnTyrIleLeuAlaSerSerAlaSer 225
QY 241 -----TACTATCCAGTCAACGAGCGGTTTCCGCGGATCACTATGCG 282
Db 226 AlaGlyGluThrProTyrPheTyrThrSerAsn-----IleProProAsnGluSerAla 253
QY 283 GTG-----ATGATCTCCAAGACGCGCACGAACCGCGGAC 318
Db 254 ValAlaSerValIleMetLeuGlnSerMetPheAsnSerThrGlyAsnTyrLysValGlu 273
QY 319 TTCACGCTGTTTTCGAAACCGCTTAACGCGAATAATAAGGCG----- 363
Db 274 PheThrIleAsnTyr-----IleProAsnGlyProAsnSerIleuValValTyrLeu 291
QY 364 -----GGAGCAAGATTGCT---CTTCCACGAGAACCAAT 396
Db 292 SerAspLeuTyrPheGluPheLeuGlySerArgTyrGlyValLeuGlyThrAspAsn 311
QY 397 GCGGCC 402
Db 312 GlyAla 313

RESULT 13
G69189
hypothetical protein MTH674 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: G69189
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T
; Qiu, D.; Spadafora, R.; Viciore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
K.I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

19	72	9.8	307	7	BLAC_MYCTU	Q10670 mycobacteri
20	72	9.8	519	1	VP6_MTV	P12325 wound tumor
21	72	9.8	1953	1	BIGA_SALTY	P25927 salmonella
22	71.5	9.8	474	1	SHU1_ECOLI	P09745 escherichia
23	71	9.7	459	1	HENO_PIG	P50828 sus scrofa
24	71	9.7	505	1	GUNN_ERMCA	Q47096 ewingia car
25	71	9.7	519	1	VP6_MTVNJ	P31610 wound tumor
26	71	10.1	2404	1	SOM_MOUSE	Q09447 mus musculus
27	70.5	9.6	607	1	XYNN_INOPIA	P29127 neocallimast
28	70	9.6	221	1	GUNN_EMEINI	P55323 emericella
29	70	9.6	1045	1	XUNB_CELFI	P26325 cellulomonas
30	70	9.6	1049	1	ITAS_HUMAN	P08648 homo sapien
31	69.5	9.5	312	1	PGTE_SALTY	P06185 salmonella
32	69.5	9.5	547	1	FUMA_ECOLI	P00923 escherichia
33	69.5	9.5	579	1	FUMA_SALTY	P40720 salmonella
34	69.5	9.5	1044	1	ITAY_MOUSE	P43406 mus musculus
35	69.5	9.5	1176	1	CIAG_BACTU	P08515 bacillus th
36	69	9.4	249	1	GEP1_YEAST	P41777 saccharomyco
37	69	9.4	463	1	STHA_PSEAE	P08515 bacillus th
38	69	9.4	517	1	STHA_PSEAE	P57112 pseudomonas
39	69	9.4	517	1	CPNI_RANCA	P52104 rna catestp
40	69	9.4	705	1	CMBA_BACSU	Q02113 bacillus su
41	68.5	9.4	958	1	TRH_DROME	Q02113 drosophila
42	68.5	9.4	321	1	CYSJ_HOMAM	P25784 homarus ame
43	68.5	9.4	521	1	NPRE_BACSU	P06142 bacillus su
44	68.5	9.8	1029	1	ITAY_HUMAN	P06576 homo sapien
45	68.5	9.8	1229	1	NI21_HUMAN	Q09213 homo sapien
					POL1_RTRVP	P27502 rice tungrei

Fletcher H.M., Schenkwein H.A., Macrina F.L.;
 "Cloning and characterization of a new protease gene (prtc) from
 Porphyromonas gingivalis".
 Infect. Immun. 62:4279-4286(1994).
 [2]
 ERRATUM.
 Fletcher H.M., Schenkwein H.A., Macrina F.L.;
 Infect. Immun. 62:5707-5707(1994).
 -I- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENABLE
 P. GINGIVALIS TO EVADE COMPLEMENT-MEDIATED KILLING DURING THE
 IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS
 AND IS A VIRULENCE FACTOR.
 -I- SUBCELLULAR LOCATION: IN MEMBRANE VESICLES.
 -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.

 THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON
 USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announcements>
 OR SEND AN EMAIL TO license@isb-sib.ch).
 EMBL, L27483; AAAA1298.1; -.
 ISSP; P23882; IEMT.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Landsberg erecta;
 RX MEDLINE=94089686; Pubmed=7505443;
 RA Corey E.J., Matsuda S.P.T., Bartel B.;
 RT "Isolation of an Arabidopsis thaliana gene encoding cycloartenol
 RT synthase by functional expression in a yeast mutant lacking
 RT lanosterol synthase by the use of a chromatographic screen";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11628-11632(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Columbia;
 RX MEDLINE=20083487; Pubmed=10617197;
 RA Liu X., Kaul S., Rounsley S.D., Shea J.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.D., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:761-768(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Columbia;
 RX Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC)";
 RL Submitted (Apr-2002) to the EMBL/GenBank/DDI databases.
 CC -1- CATALYTIC ACTIVITY: (S)-2,3-epoxysqualene -> cycloartenol.
 CC -1- SIMILARITY: BELONGS TO THE TERPENE CYCLASE/MUTASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 PPTB REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U02555; AAC04931.1; -
 DR EMBL: AC005171; AAM15013.1; -
 DR EMBL: AY094394; AAM19773.1; -
 DR InterPro: IPR001330; PrenylTrans.
 DR InterPro: IPR002365; Terpene_synth.
 DR Pfam: PF00432; prenyltrns; 3.
 DR PROSITE: PS01074; TERPENE_SYNTASES; 1.
 KW Isomerase; Repeat.
 FT REPEAT 147 188 PPTB 1.
 FT REPEAT 512 557 PPTB 2.
 FT REPEAT 589 629 PPTB 3.
 FT REPEAT 638 679 PPTB 4.
 FT REPEAT 700 741 PPTB 5.
 FT CONFLICT 501 501 A -> E (IN REF. 1).
 SQ SEQUENCE 759 AA; 86033 MW; DBA75CC57BB1F74D CRC64;
 Alignment Scores:
 Pred. No.: 1.3 length: 759
 Score: 80.00 Matches: 25
 Percent Similarity: 43.688 Conservative: 13
 Best Local Similarity: 28.748 Mismatches: 15
 Query Match: 10.948 Indels: 34
 DB: 1 Gaps: 7
 US-09-980-370-5 (1-402) x CAS1_ARATH (1-759)
 QY 37 GGAGGCGACGCGGAGGAACTGACTATGATCGGATCGGATGGAGGTTGGCTC 96
 DB 188 GlycylglycylPro-----AsnAspClyAspClyAsp----- 197

QY 97 TGCTGCTCCGACAAATTCGACTGCTACAGCTCATGCGCGC---ACCAACGTAGTA 153
 DB 198 ---MetClnuyslslyArg---AsprtleLeuasnHslslyAlaIatrnAsnIlethr 215
 QY 154 AGC-----TCTTCTCATGAGTGGN---ATGGCT 180
 DB 216 SerTrpGlyLysMetTrpLeuSerValLeuGlyAlaIlePheGlyLysAsnAsnPro 235
 QY 181 TTGAATCTGATTAATCATTCATCTCAAGATGTACAGCGCGACCAAGTAAGTAC 240
 DB 236 LeuProGlnIleTrpLeuLeu-----ProTyr 245
 QY 241 TACTATCCAGTCAACGACGGT 261
 DB 246 PheLeuProIleHisProGly 252
 RESULT 4
 THER_BACTH STANDARD: PRT; 316 AA.
 AC P00800;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Thermolysin (EC 3.4.24.27).
 OS Bacillus thermoproteolyticus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1427;
 RN [1]
 RP SEQUENCE.
 RA Titani K., Hermodson M.A., Ericsson L.H., Walsh K.A., Neurath H.;
 RT "Amino-acid sequence of thermolysin";
 RL Nature New Biol. 238:35-37(1972).
 RN [2]
 RP ACTIVE SITE.
 RX MEDLINE=74052951; Pubmed=4808703;
 RA Burshtein Y., Walsh K.A., Neurath H.;
 RT "Evidence of an essential histidine residue in thermolysin";
 RL Biochemistry 13:205-210(1974).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=83085812; Pubmed=7175940;
 RA Holmes M.A., Matthews B.W.;
 RT "Structure of thermolysin refined at 1.6-A resolution";
 RL J. Mol. Biol. 160:623-639(1982).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=75041142; Pubmed=4214815;
 RA Matthews B.W., Weaver L.H., Kester W.R.;
 RT "The conformation of thermolysin";
 RL J. Biol. Chem. 249:8030-8044(1974).
 RN [5]
 RP STRUCTURE BY NMR OF 255-316.
 RX MEDLINE=95086079; Pubmed=7993910;
 RA Rico M., Jimenez M.A., Gonzalez C., de Filippis V., Fontana A.;
 RT "NMR solution structure of the C-terminal fragment 255-316 of
 RT thermolysin: a dimer formed by subunits having the native
 RT structure";
 RL Biochemistry 33:14834-14847(1994).
 RN [6]
 RP STRUCTURE BY NMR OF 205-316.
 RX MEDLINE=97452621; Pubmed=9305992;
 RA Conejero-Lara F., Gonzalez C., Jimenez M.A., Padmanabhan S.,
 RA Mateo P.L., Rico M.;
 RT "NMR solution structure of the 205-316 C-terminal fragment of
 RT thermolysin. An example of dimerization coupled to partial
 RT unfolding";
 RL Biochemistry 36:11975-11983(1997).
 CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Xaa-[1-Leu > Xaa-1]-Phe.
 CC -1- COPACOPOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS WHICH
 CC CONTRIBUTE TO THE UNUSUAL THERMAL STABILITY OF THIS ENDOPEPTIDASE.
 CC TWO ARE CLOSE TOGETHER AND PARTICIPATE IN A NETWORK OF SALT LINKS
 CC ALSO INVOLVING ASP-136, GLU-177, LYS-182, ASP-185, GLU-190, AND

FT	HELIX	234	246	
FT <td>STRAND</td> <td>248</td> <td>250</td> <td></td>	STRAND	248	250	
FT	TURN	251	252	
FT	STRAND	253	255	
FT	HELIX	260	273	
FT	TURN	274	274	
FT	TURN	277	278	
FT	HELIX	281	296	
FT	TURN	298	299	
FT	HELIX	301	312	
FT	TURN	313	314	
SO	SEQUENCE	316 AA;	34333 MM;	004EDAF478744BB4 CRC64;
Alignment Scores:				
Pred. No.:		1.3	Length:	316
Score:		79.50	Matches:	40
Percent Similarity:		41.98%	Conservative:	28
Best Local Similarity:		24.69%	Mismatches:	53
Query Match:		10.88%	Indels:	41
Dbs:		1	Gaps:	10
US-09-980-370-5 (1-402) x THER_BACSTH (1-316)				
OY	25	TCCTCTACAT	-----GGAGAGGACCGAGGATGGACT-----	ACTATC 66
Db	102	SerSerValHisIysIysSerIcngIyUryrAsnAlaIlePrrAsnGlySerIcIuMetVal	121	
OY	67	GATGCCCATGGCGAGTGTGAGGGTGGCTCTGTCTGTCTTCGGACAAATGGACTGGCTC	126	
Db	122	TyTcIyAspIyIyAspIyIyGlnIhThrPheIleProIeuSerIyGly---	139	
OY	127	ACAGCTCATGGCGGACCAAGCTAGTAAGCTCTTTCATGAATGGAATGGCTTTCAT	166	
Db	140	ValAlaHisIcIuIeuThrHisAlaValThrAspIyTrHr---	158	
OY	187	CCGATATACATCTCTCATCTCAAG-----GATTTACAGCGGACAGAGTAAAG	237	
Db	159	AsnIuSerIyAlaIleAsnGlnAlaIleSerAspIlePheGly---	177	
OY	238	TACTACTATTCACACTCAAGCAGCGTTT-----	270	
Db	178	PheTyTAlaIleAsnIyAsnProAspTrpGlnIleGlyGlnAspValTyTrHrProGlyIle	197	
OY	271	-----	279	
Db	198	SerGlyAspSerIeuIySerMetSerAspProAlaIyTyTcIyAspProAspHisTyTc	217	
OY	280	GCGGTGATGATCTCCCAAGCGGCGACGAAGCCGGAGACTTC-----	333	
Db	218	SerIyAspIyTrHrThyIyTrHrGlnAspAsnGlyGlyValIhIleAsnSerGlyIleIle	237	
OY	334	GAGAAACCGCTTACCGAATAATAGAGCGGCGAGCAATTCGGTCTTTCACCGAAGCC	353	
Db	238	AsnIyAlaIleAlaIyTrIleuIleSerGlnGlyTrHrIstTyTcIyValSerValIcIy	257	
OY	394	AATGGC 399		
Db	258	IleGly 259		
RESULT 5				
NPRS_BACST		STANDARD:	PRT:	551 AA.
AC	P41133:			
DT	01-NOV-1995 (Rel. 32, Created).			
DT	01-NOV-1995 (Rel. 32, Last sequence update).			
DT	01-OCT-1996 (Rel. 34, Last annotation update).			
DE	BacilloIySln precursor (EC 3.4.24.28) (Neutral protease).			
GN	NPRS OR NPRM.			
OS	Bacillus stearothermophilus.			
CC	Bacteria; Firmicutes; Bacillales; Geobacillus.			
NCBI	NCBI:taxID:1422;			
SP	SEQUENCE FROM N.A., AND SEQUENCE OF 236-239.			


```
DR InterPro: IPRO01823; Ald1-epimerase.  
DR Pfam: PF01263; Aldose_aldose_epim: 1.  
DR ProDom: PD006407; Ald1-epimerase: 1.  
DR PROSITE: PS00545; ALDOSE_1_EPIMERASE: 1.  
KW Isomerase.  
FT ACN SITE 180 180 POTENTIAL.  
SQ SEQUENCE 348 AA: 39012 MW: 17C3E027ABEF92 CRC64:  
  
Alignment Scores:  
Pred. No.: 1.69 Length: 348  
Score: 78.50 Matches: 26  
Percent Similarity: 42.86% Conservative: 31  
Best Local Similarity: 30.95% Mismatches: 10  
Query Match: 10.74% Indels: 17  
DB: Gaps: 3  
  
US-09-980-370-5 (1-402) x GAA_M_STRPTR (1-348)  
QY 85 GAGGGTGGCTGTCTGTCCTCCGACAAATTGCAGCTCACAGCTCATGCCGACC 144  
Db 99 GLNGLYASPAANCYSLEU-----HSLGYLGYPCO 108  
QY 145 AACGTGAATGACTTTTCATGCAATGTAATGCGTTTGATCTCTGATCTATCTCATC 204  
Db 109 LYGSLYMELTNVALDINSPTIRPSNTYRVALTHIRASLENLSAASAPSYTRVALGL 128  
QY 205 TCAAAGSANTCTTACAGCCCAACGAAGTAAAGTACTACTTCCATGCACAGCGTTTT 264  
Db 129 ThrLLP-----PhellelrgatgLeuYrSerValIaspGlyPhe 142  
QY 265 CCCGGGATCATCTTCGCGTGATGATCTCCAAAGAGGCGACAGCGCGGAGACTTCACG 324  
Db 143 PGGILYASP---VALThrfalserileSerTrfargLeuAsnAsnAsnArgLeuThr 161  
QY 325 GTTGTTTTGCA 336  
Db 162 ILeuapheglu 165  
  
RESULT 7  
TNAL_SYSTM STANDARD: PRT: 453 AA.  
ID TNAL_SYSTM  
AC P31014;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tryptophanase 1 (EC 4.1.99.1) (L-tryptophan indole-lyase 1) (Tnase 1).  
GN TNAL  
OS Symbiodacterium thermophilum.  
OC Bacteria: Actinobacteria: Actinobacteria (class); Symbiodacterium.  
OX NCBI_TaxID=2734.  
RN (1)  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-36.  
RC STRAIN=IAM 13621;  
RX MEDLINE=92384583; PubMed=1339259;  
RA Hirahara T., Suzuki S., Horiouchi S., Beppu T.;  
RT "Cloning, nucleotide sequences, and overexpression in Escherichia coli of tandem copies of a tryptophanase gene in an obligately symbiotic thermophile, Symbiodacterium thermophilum.";  
RL Appl. Environ. Microbiol. 58:2633-2642(1992).  
RL CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> indole + pyruvate + NH(3).  
CC -|- COPFACTOR: Pyridoxal phosphate.  
CC -|- PATHWAY: Tryptophan catabolism.  
CC -|- SUBUNIT: Homotrimer (by similarity).  
CC -|- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement_e or send an email to license@isb-sib.ch).
```

```

CC      EMBL: AB010832; BAA24688.1; -.
DR      PTR: A49022; A49022.
DR      HSSP: P31013; IPR1.
DR      InterPro: IPR001597; Beta_elim_lyase.
DR      Pfam: PF01212; Beta_elim_lyase; 1.
DR      ProDom: PD005927; Beta_elim_lyase; 1.
DR      ProSite: PS00853; BETA_ELIM_LYASE; 1.
KW      Tryptophan catabolism; Lyase; Pyridoxal phosphate.
FT      BINDING 254 254 PRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ      SEQUENCE 453 AA: 51012 MW: CADI332655C08812 CRC64.

Alignment Scores:
pred. No.:      1 76      Length:      453
Score: No.:      78.50      Matches:      25
Percent Similarity: 42.55%      Conservative: 15
Best Local Similarity: 26.60%      Mismatches: 41
Query Match:      10.74%      Indels: 13
DB:              1      Gaps:      4

US-09-980-370-5 (1-402) x TNA1_SYMTH (1-453)
OY      115 TTGGACGCGCTGACAGCTCATGCGGCGACCAACGTAAGCTGCTTCTCATGATGCA 174
DB      44 ITLSPLEUETH---ASPSEGLYTHGLYALAMEUSEGLINPHEGLINPSEALA 62
OY      175 ATGGCTTTGAATCTGATACATATCTC-----ATCTCAAGCAT 213
DB      63 MEEHLEUdLYASPISerTYrAlaGLyAlaSerSerTYrTYrAlaGLyLeuLYSGluThr 82
OY      214 GTTACAGCGCAGCAGAGAGTAAGTACTACTATCTTCAGTCAACAGCAGCGTTTCCCGGCGAT 273
DB      83 ValThrAspLeuThrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 102
OY      274 CACTATGCGGCTGATGATCTTCGAAAGACGGGCGACGCGGAGAGACTC-----ACG 324
DB      103 -----LysValAlaPheSerGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
OY      325 GTTGTTCGAGAGAACGCGCTTAACGGAATTAATAAGCGCGCA 366
DB      121 MetPhePheAspThrThrArgGlyLysValGlnLeuGlyGly 134

RESULT 8
GUNZ_ERWCH STANDARD: PRT: 426 AA.
ID      GUNZ_ERWCH
AC      P07103;
DT      01-APR-1988 (Rel. 07, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Endoglucanase Z precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase Z)
DE      (Cellulase Z) (EGZ).
OS      CELZ OR CELS.
OS      Erwinia chrysanthemi.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Pectobacterium.
OX      NCBI_TaxID:556;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN:3937.
RX      MEDLINE:88216177; PubMed:2835589;
RX      Guisepi A., Cami B., Aymeric J.-L., Ball G., Creuzet N.;
RT      "Homology between endoglucanase Z of Erwinia chrysanthemi and
RT      endoglucanases of Bacillus subtilis and alkalophilic Bacillus.";
RL      Mol. Microbiol. 2:159-164(1988).
RN      [2]
RP      REVISIONS, AND DISULFIDE BOND.
RC      STRAIN:3937.
RX      MEDLINE:84203057; PubMed:8152378;
RX      Bortoli-german I., Brun E., Py B., Chipaux M., Barris F.;
RT      "Periplasmic disulphide bond formation is essential for cellulase
RT      secretion by the plant pathogen Erwinia chrysanthemi.";
RL      Mol. Microbiol. 11:545-553(1994).
RN      [3]

```

Alignment Scores:		Pred. No.:	1,97	Length:	426
Score:	78.00	Matches:	30		
Percent Similarity:	44.55%	Conservative:	15		
Best Local Similarity:	29.70%	Mismatches:	24		
Query Match:	10.67%	Indels:	32		
DB:	1	Gaps:	7		
US-09-980-370-5 (1-402) x GUNZ_ERMCH (1-426)					
OY	7	TTACAGGAAACGTTGAGCTCTTACTCATGAGAGGCA-----	45		
Db	232	TYRTHLEUHIHPHERYVALAOLYTHRHISGLISLUSELEALAGSALYALALAGIN	251		
OY	46	-----CCAGCGAATGAGCATACATGATGCGCATGCG	78		
Db	252	ALALEUASNSNGLYLLEALALEUPHEVALTHGLUTRPGLYTHRVALASNLALASPLY	271		
OY	79	GATGGTAGGAGTGGCTGCTGTCGTCCGAGCAATGTGAC---TGSCCTACAGCTCAT	135		
Db	272	ASNGLYGLY-----VALASNGLINTHGLIETHASALATTPVALTRPHENET	287		
OY	136	GGCGGACACACAGTAAAGCTCTTTCATGCAATGAGATGCGCTTGATCCTGATAC	195		
Db	288	ARGSPASNSN---ILLESASNLASLRTP-----ALALEUASNSPLYSASN	303		
OY	196	-----TATCTCATCTCAAGATGTTACA---GGCGACAGCAAGCTA	234		
Db	304	GLUGLYLASERTHRYTRYTPROASPSELYASNLLEUTHRGLUSERCYLYLSVAL	323		
OY	235	AAG 237			
Db	324	Lys 324			
RESULT 9					
NP_RE_BACBR	ID	NP_RE_BACBR	STANDARD:	PRT:	527 AA.
AC	P43263:				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	01-NOV-1995 (Rel. 32, Last annotation update)				
DE	Bacillolysin precursor (EC 3.4.24.28) (Neutral protease).				
GN	NR.				
OS	Bacillus brevis.				
OC	Bacteria: Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.				
OX	NCBI_TaxID=1393;				
NP	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-7882;				
RX	MEDLINE-91148616; Pubmed-2290427;				
RA	AVAKOV A.S., BOLOTOV A.P., SOROKIN A.V.;				
RL	"Structure of the Bacillus brevis metalloprotease gene.";				
RP	Mol. Biol. (Mosk). 24:1363-1372(1990).				
RP	SEQUENCE OF 224-228, AND CHARACTERIZATION.				
RX	MEDLINE-91148616; Pubmed-2127074;				
RA	KALIDLOVA N.V., AKIMKINA T.V., KHODOVA O.D., KOSTOV S.V.;				
RA	Strongin A.V.;				
RT	Analysis of the structure of Bacillus brevis neutral proteinase and				
RT	its biosynthesis in Bacillus subtilis cells.";				
RL	Mol. Biol. (Mosk). 24:1381-1392(1990).				
CC	-1- FUNCTION: THERMOLABILE EXTRACELLULAR ZINC METALLOPROTEASE.				
CC	-1- CATALYTIC ACTIVITY: Similar, but not identical, to that of				
CC	thermolysin.				
CC	-1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY				
CC	SIMILARITY).				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				

Db 84 ValaspaValThrpHeserGlyThrProAspGlyVal 95

RESULT 11

ITAV_CHICK STANDARD: PRT: 1034 AA.

ID ITAV_CHICK STANDARD: PRT: 1034 AA.

AC P26008;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Integrin alpha-V precursor (Vitronectin receptor alpha subunit).

GN ITGAV.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA MEDLINE=91104936; PubMed=1703004;

RA Bossy B., Reichardt L.F.;

RT "Chick integrin alpha V subunit molecular analysis reveals high conservation of structural domains and association with multiple beta subunits in embryo fibroblasts."

RT Biochemistry 29:10191-10198(1990).

RL -1- FUNCTION: THE ALPHA-V INTEGRINS ARE RECEPTORS FOR VITRONECTIN, CYTOTACTIN, FIBRONECTIN, FIBRINOGEN, LAMININ, MATRIX METALLOPROTEINASE-2, OSTEOPOINTIN, PROTHROMBIN, THROMBOSPONDIN AND VON WILLEBRAND FACTOR. THEY RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A DISULFIDE BOND. ALPHA-V ASSOCIATES WITH EITHER BETA-1, BETA-3, BETA-5, BETA-6 OR BETA-8.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: M60517; AAA49138.1; -

DR PIR: A36108; A36108.

DR HSSP: P06756; IJV2.

DR InterPro: IPR000413; Integrin_alpha.

DR Pfam: PF01839; FG-GAP: 5.

DR PRINTS: PRO1185; INTEGRIN.A.

DR SMART: SM0191; InL_alpha: 5.

DR PROSITE: PS00242; INTEGRIN_ALPHA: 1.

DR Integrin: Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Calcium.

KW SIGNAL: Repeat; Calcium.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 1034 INTEGRIN ALPHA-V.

FT CHAIN 20 875 INTEGRIN ALPHA-V HEAVY CHAIN (POTENTIAL).

FT CHAIN 877 1034 INTEGRIN ALPHA-V LIGHT CHAIN (POTENTIAL).

FT DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 979 1002 POTENTIAL.

FT DOMAIN 1003 1034 CYTOPLASMIC (POTENTIAL).

FT REPEAT 35 96 FG-GAP 1.

FT REPEAT 101 166 FG-GAP 2.

FT REPEAT 167 224 FG-GAP 3.

FT REPEAT 237 290 FG-GAP 4.

FT REPEAT 291 350 FG-GAP 5.

FT REPEAT 355 415 FG-GAP 6.

FT REPEAT 420 473 FG-GAP 7.

FT CA_BIND 248 256 POTENTIAL.

FT CA_BIND 302 310 POTENTIAL.

FT CA_BIND 367 375 POTENTIAL.

FT CA_BIND 431 439 POTENTIAL.

FT SITE 1005 1009 GEFKR MOTIF.

FT DISULFID 77 85 BY SIMILARITY.

FT DISULFID 126 146 BY SIMILARITY.

FT DISULFID 160 173 BY SIMILARITY.

FT DISULFID 479 488 BY SIMILARITY.

FT DISULFID 494 551 BY SIMILARITY.

FT DISULFID 612 618 BY SIMILARITY.

FT DISULFID 684 697 BY SIMILARITY.

FT DISULFID 838 890 INTERCHAIN (BY SIMILARITY).

FT DISULFID 895 900 BY SIMILARITY.

FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 690 690 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 860 860 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 931 931 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 966 966 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 1034 AA; 114388 MW; D76B08BA692DC684 CRC64;

Alignment Scores:

Pred. No.: 4.14 Length: 1034

Score: 75.50 Matches: 36

Percent Similarity: 37.75% Conservative: 21

Best Local Similarity: 23.84% Mismatches: 47

Query Match: 10.33% Indels: 47

DB: 1 Gaps: 9

US-09-980-370-5 (1-402) x ITAV_CHICK (1-1034)

OY 19 TTGAGCTCTTCTACTCATGGAGAG---GCACGACGGAATGACTACTATGATCCGCAT 75

Db 149 Pheaspplserlysservalglutrylarpocysargserthrtrthleasplalasp 168

OY 76 GCGCATGTGAGGGGTGGCTGTCTCTCCGACAATGAGCTGCTCAGACATCAT 135

Db 169 Glylnlly-----Phecysglnglylgheserleasphethllys----- 183

OY 136 GCGGCGACACAGTAGTA-----AGCTTTTTCATGGAATGGAATGGCTTTTG 183

Db 184 ---GlyasparyallleuGllylproglyserPheyrtrpglnglylnleuile 202

OY 184 AATGCT----- 189

Db 203 SerasparyalalagluilleleualalystyraspserlysvaItyrserthlystyr 222

OY 190 GATACATATCTCATCTCAAGAGATGTACAGGCGCAGGAAGTAAGTACTACTATCA 249

Db 223 Aspaspllnleualarhrarproalaserlala----- 234

OY 250 GTCAACGACGGTTTCCCGGGGATCACTATGCGGTGATGATCTTCAAGACGGCGACG 309

Db 235 PheaspserlyrleuGly-----TyrserValalaval----- 246

OY 310 GCGGACACTTCACG---GTTGTTTCGAGGAAGAACGCTTAACGGAATTAATAGCGCGCA 366

Db 247 ---GlyaspheaserlyaspGlyllegluasphevalserglyvalProalgalalala 265

OY 367 GCAAGATTGCTCTTTCACGGAAGCAATGCG 399

Db 266 ArgThrleuGlymetValserIleTyraSngly 276

RESULT 12

Ther_BACST STANDARD: PRT: 548 AA.

DE Putative thiosulfate sulfurtransferase SSEA (EC 2.8.1.1).
GN SSEA OR RV3283 OR MT3382 OR MTCY71.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_taxid=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Blahai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.
CC -1- SIMILARITY: BELONGS TO THE RHODANSE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 292771; CAB07066.1;
CC EMBL: AE007147; AAK47725.1;
CC HSRP: P52197; 1E0C.
DR TIGR: MT3382;
DR TubercuList: RV3283;
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR001307; Rhodanese.
DR Pfam: PF00581; Rhodanese; 2.
DR SMART: SM00450; RHOD. 2.
DR PROSITE: PS00380; RHODANSE_1; 1.
DR PROSITE: PS00683; RHODANSE_2; 1.
KW Hypothetical protein; transferase; Complete proteome.
FT ACT_SITE 245 245 BY SIMILARITY.
SQ SEQUENCE 297 AA: 33320 MW: 5930ABSP9C63A533 CRC64;

Alignment Scores:
Pred. No.: 5.7 Length: 297
Score: 73.50 Matches: 22
Percent Similarity: 45.35% Conservative: 17
Best Local Similarity: 25.58% Mismatches: 30
Query Match: 10.05% Indels: 17
DB: 1 Gaps: 4

US-09-980-370-5 (1-402) x THF2_MVCTU (1-297)
QY 118 GACTGCTCAGACCTCATGCGGACCAAGTAGTAAGCTCTTCTCATGAGATGATG 177
DB 24 AsprtrpleuseralshsmetglyAla-----ProGlyLeu 35
QY 178 GCCTTTC---AATCTCATATCATCTCATCTCAAGATGTT-----ACAGGCGCA 225

DB 36 AlaIleValGluserIsprgluAspValLeuLeuTyrAspValGlyHisIleProGlyAla 55
QY 226 ACGAAGCTAAAGTACTACTATCCAGTCAACGACGCT-----TTTCCCGGG 270
DB 56 ValIysIleAsprtrphIsTrAspLeuAsnAspProArGValArGAspTyrIleAsnGly 75
QY 271 GATCAGTATGCGGTGATGATCTCCCAAGACGGGACGAACCGCGGAGACTTCACGCTGT 330
DB 76 GluGlnPheAlaGluLeuMetAspArgLysGlyIleAlaArgAspAspThrValValIle 95
QY 331 TTCGAAGAAACCGCTTAC 348
DB 96 TyrGlyAspIysSerAsn 101

Search completed: June 26, 2003, 00:44:53
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: June 25, 2003, 23:06:58 : Search time 35.5 Seconds
(Without alignments)
4666.532 Million cell updates/sec

Title: US-09-980-370-5

Perfect score: 731
Sequence: 1 gcaagcttcacggaacglt.....ccacggaagcgaagcgcgc 402

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlh
-O=/gen2.1/USPTO.spool/US09980370/runet.19062003.174442.8856/app_query.fasta.1.583
-DB=SPRMBL_21 -QWTF=fastan -SUFFIX=trspc -MINMATCH=0.1 -LOOPECI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09980370.ecgn.1.1.57.grunt.19062003.174442.8856 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	716	97.9	1223	2	Q9ZNB5 porphyromon

2	716	97.9	1687	2	Q9R9B7	Q9R9B7 porphyromon
3	716	97.9	1704	2	Q51816	Q51816 porphyromon
4	715	97.8	1097	2	P72196	P72196 porphyromon
5	715	97.8	1732	2	O52050	O52050 porphyromon
6	715	97.8	1732	2	O51817	O51817 porphyromon
7	715	97.8	1732	2	O51817	O51817 porphyromon
8	714	97.7	1706	2	O07442	O07442 porphyromon
9	711	97.3	1723	2	O51839	O51839 porphyromon
10	709	97.0	1358	2	P72197	P72197 porphyromon
11	709	97.0	1723	2	P69667	P69667 porphyromon
12	706	96.6	1706	2	P72194	P72194 porphyromon
13	705	96.6	1706	2	O51838	O51838 porphyromon
14	205	28.0	925	2	O9F4J0	O9F4J0 porphyromon
15	180	24.6	312	2	O9K1B3	O9K1B3 porphyromon
16	90.5	12.4	218	16	O9J215	O9J215 neisseria m
17	89	12.2	286	17	O9V211	O9V211 pyrococcus
18	89	12.2	418	9	O80060	O80060 staphylococ
19	89	12.2	1225	9	O8SDK2	O8SDK2 staphylococ
20	88.5	12.1	1261	9	O9MBN9	O9MBN9 staphylococ
21	88.5	12.1	218	16	O9J010	O9J010 neisseria m
22	87	11.9	910	3	P87211	P87211 orpiniomyces
23	85.5	11.7	2275	17	O8TSE8	O8TSE8 methanocarc
24	84	11.5	494	2	O86099	O86099 alteromonas
25	83.5	11.4	256	17	O8TPD2	O8TPD2 methanocarc
26	83.5	11.4	548	2	O45779	O45779 bacillus th
27	83.5	11.4	603	10	O9M653	O9M653 polyanatum
28	83	11.4	373	16	O53926	O53926 mycobacteri
29	83	11.4	806	3	O9P478	O9P478 agaricus bl
30	82.5	11.3	536	17	O97YP2	O97YP2 sulfolobus
31	81.5	11.1	966	17	O26770	O26770 methanobact
32	81	11.5	355	10	O9AUL4	O9AUL4 oryza sativ
33	81	11.1	1372	2	O54151	O54151 shigella fl
34	81	11.1	1373	2	O9A158	O9A158 shigella fl
35	79.5	10.9	534	5	O25402	O25402 lymanaea sta
36	79.5	10.9	735	16	O8XNP9	O8XNP9 clostridium
37	78.5	10.7	348	2	O8VS90	O8VS90 streptococc
38	78.5	10.7	449	2	O8VS84	O8VS84 streptococc
39	78	10.7	499	2	O45532	O45532 bacillus su
40	78	10.7	785	3	O9P8H5	O9P8H5 humicola in
41	78	10.7	803	10	O8W075	O8W075 sorghum bic
42	78	10.7	1606	12	O931D6	O931D6 white spot
43	77.5	10.6	438	3	O9P8K6	O9P8K6 leptosphaer
44	77.5	10.6	1039	5	O9VQB1	O9VQB1 drosophila
45	77.5	11.0	3014	12	O39928	O39928 hepatitis c

ALIGNMENTS

RESULT 1

ID Q9ZNB5 PRELIMINARY: PRT: 1223 AA.

AC Q9ZNB5: 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 130K-HMGD (Fragment)
OS Porphyromonas gingivalis (Bacteroides gingivalis)
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RX MEDLINE=99143166; PubMed=9988746;
RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.;
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products."
RT J. Biol. Chem. 274:5012-5020(1999).
KL EMBL: AB019363; BAA34341.1;
DR InterPro: IPR000977; DNA ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.

DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT NON_TER 1
 SO SEQUENCE 1223 AA: 131542 MW: 00225CD2BA9F1B3 CRC64:

Alignment Scores:
 Pred. No.: 2,48e-62 Length: 1223
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: Gaps: 0

US-09-980-370-5 (1-402) x Q92NB5 (1-1223)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGCAGCAGCGGAATGACT 60
 DB 648 AlaasphenethrgrluthrphengluserthrhisglualaProalaglutrpthr 667
 QY 61 ACTATCGATGCCGATGGCGATGGTGGCTCTGCTCTCTTCGCGACATTTGGAC 120
 DB 668 Thrileaspaalaspglaspilglnlytrpleucysleuserserglglnleuasp 687
 QY 121 TGGCTCACAGCTCATGGCGGCGACCAAGTACGTTCTTCATGATGATGATGCT 180
 DB 688 Trpleuthralahisgllythrnsvalvalalaserphesertrrpsnnglymetala 707
 QY 181 TTGAATCCTGATACATCTCATCTCAAGATGTTACAGCGGCAAGAGTAAGTAC 240
 DB 708 LeuasnProaspasnlyrleuileserlyssaspvalthrglYalathrlystlyr 727
 QY 241 TACTATCCAGTCAACGAGCGTTTCCCGGATCAGTACGCGTATGATCTCCAGACG 300
 DB 728 TyrtyralaValaAsnaaspilypheproglaspshistylalaValmetileserlysthr 747
 QY 301 GGCACGAGCGCGGAGACTTTCACGCGTGTTCGAGAGAAAGCGCTTAACGAATAAG 360
 DB 748 GlythrnsnalaglasphepthervalValpneglugluthrProasnnglylleasnlys 767

OY 361 GCGGAGCAGCAATTCGCTTTTCACGAGGAAGCAATGGCGCC 402
 DB 768 Glyglylaalargpneclyleuserthrglualaasnnglyala 781

RESULT 2

Q9R9B7 PRELIMINARY: PRT: 1687 AA.

ID Q9R9B7
 AC 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hemagglutinin/protease.
 GN HAGE.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria: CFb group: Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=381.
 RA Han N., Dong H., Prognulke-Fox A.;
 RT "Cloning and characterization of hage from P. gingivalis 381."
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF026946; AAD01810.1; -
 DR HSSP: P95493; 1CVR.
 DR MEROPS: C25.001; -
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW PROTEINASE.
 SO SEQUENCE 1687 AA: 183702 MW: D085B516A39FE70 CRC64:

Alignment Scores: 2.61e-62 Length: 1687
 Pred. No.: 2,61e-62 Length: 1687

Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: Gaps: 0

US-09-980-370-5 (1-402) x Q9R9B7 (1-1687)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGCAGCAGCGGAATGACT 60
 DB 1120 AlaasphenethrgrluthrphengluserthrhisglglnalProalaglutrpthr 1139
 QY 61 ACTATCGATGCCGATGGCGATGGTGGCTCTGCTCTCTTCGCGACATTTGGAC 120
 DB 1140 Thrileaspaalaspglaspilglnlytrpleucysleuserserglglnleuasp 1159
 QY 121 TGGCTCACAGCTCATGGCGGCGACCAAGTACGTTCTTCATGATGATGATGCT 180
 DB 1160 Trpleuthralahisgllythrnsvalvalalaserphesertrrpsnnglymetala 1179
 QY 181 TTGAATCCTGATACATCTCATCTCAAGATGTTACAGCGGCAAGAGTAAGTAC 240
 DB 1180 LeuasnProaspasnlyrleuileserlyssaspvalthrglYalathrlystlyr 1199
 QY 241 TACTATCCAGTCAACGAGCGTTTCCCGGATCAGTACGCGTATGATCTCCAGACG 300
 DB 1200 TyrtyralaValaAsnaaspilypheproglaspshistylalaValmetileserlysthr 1219
 QY 301 GGCACGAGCGCGGAGACTTTCACGCGTGTTCGAGAGAAAGCGCTTAACGAATAAG 360
 DB 1220 GlythrnsnalaglasphepthervalValpneglugluthrProasnnglylleasnlys 1239
 QY 361 GCGGAGCAGCAATTCGCTTTTCACGAGGAAGCAATGGCGCC 402
 DB 1240 Glyglylaalargpneclyleuserthrglualaasnnglyala 1253

RESULT 3

Q51816 PRELIMINARY: PRT: 1704 AA.

ID Q51816
 AC 051816:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Arg-gingipain-1 proteinase.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria: CFb group: Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95138080; PubMed=7836351;
 RA Pavloff N., Potempa J., Pike R.N., Prochazka V., Kiefer M.C.,
 RA Travis J., Barr P.J.;
 RT "Molecular cloning and structural characterization of the Arg-
 gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
 proteinase-adenin polypeptide."
 RL J. Biol. Chem. 270:1007-1010(1995).
 DR EMBL: U15282; AAA69539.1; -
 DR HSSP: P95493; 1CVR.
 DR MEROPS: C25.001; -
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT CHAIN 228 719 MATURE 50-KDA CYSTEINE PROTEINASE
 FT GINGIPAIN.
 SO SEQUENCE 1704 AA: 185436 MW: 6A34B013C2A676 CRC64:

Alignment Scores: 2.62e-62 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1

```

US-09-980-370-5 (1-402) x 051816 (1-1704)
DB: 2 Gaps: 0
Query Match: 97.95% Indels: 0
OY 1 GCAGACTTCACGGGAAGCGTCGACGCTCTCATCTCATGACGAGGACACACAGCGAATGCAC7 60
Db 1137 ALAaPpethrctlnrPheluserSerThrIhISglIaIaProIaIaGlvtrPthr 1156
OY 61 ACTATCGATGCCGATGCGCATGCGATGGAGGGTTGGCTCTGCTGTCTTCGCGACAATGAC 120
Db 1157 ThrIleaaPpethrctlnrPheluserSerThrIhISglIaIaProIaIaGlvtrPthr 1176
OY 121 TGCGCTACACACTGACGGCGGCGACCAAGCAATGAGTCTTCCTCATGATGGAAAGCGT 180
Db 1177 TrpIlethrIaIaIhISglIyGlyThrAaIaValIaIaSerPheSerTrpAnISglIaIa 1136
OY 181 TTGAATCTGTAAATATCTATCTCATCTCAAGAGATGTATACAGCGCAACGAGAAAGTAC 240
Db 1197 LeuaaPpethrctlnrPheluserSerThrIleSerLysaaPpethrIhISglIaIaIa 1216
OY 241 TACTATCATGCTACCAAGCGGTTTTCCTCCGGGACACTATGCGGTGATCATCTCAAGAC 300
Db 1217 TyTtTtIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1236
OY 301 GGCAGCAAGCGCGGACATTCACGCGTGTCTTTTGAGAAACCGCTTAACGGAATAATTA 360
Db 1237 GCGTAAAGAAATGAAaPpethrctlnrPheluserSerThrIleaaPpethrIhISglIa 1256
OY 361 GCGGCAAGCAAGTATCGCTCTTCACGCAAGCAAGTACAGCGGC 402
Db 1257 GlyIyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1270
RESULT 4
AC ID P2196 PRELIMINARY: PRT: 1097 AA.
AC P72196
DT 01-FEB-1997 (TRMBLrel. 02, Created)
DT 01-FEB-1997 (TRMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TRMBLrel. 21, Last annotation update)
DE Tomb-1-linked adhesin precursor.
TLA.
GN Porphyromonas gingivalis (Bacteroides gingivalis).
NC Bacteri: CRP group, Bacteroidales; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas
OC NCBI_Taxid=837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:97386416; PubMed:9244265;
RX AA Aduse-Opoku J., Stanley K.M., Young K.A., Muir J., Rangasajan M.,
RT "The cta gene of Porphyromonas gingivalis W50, a homologue of the
RT arginine-specific porphyromonas precursor (PRPR) which shares sequence
RT similarity to Tomb-1-linked receptors."
RL J. Bacteriol. 179: 478-478(1997).
DR EMBL: Y07618; CAA68897.1;
DR Interpro: IPR009377; DNM_11998;
DR RefSeq: PF01364; RefSeq: G35;
DR Pfam: PF01364; Pfam: G35;
DR PROSITE: PS00697; DMLTIGASE_A1; UNKNOWN_1.
FT SIGNAL. 1 53 POTENTIAL.
FT SIGNAL. 1 53 POTENTIAL.
SO SEQUENCE 1097 AA; 118731 MW; 73BBA337B421F8B9 CRC64;

Alignment Scores:
Score: No: 3,076-62 Length: 1097
Percent Similarity: 99.25% Mismatches: 131
Best Local Similarity: 97.76% Conservative: 2
Query Match: 97.81% Idels: 0
Gaps: 0
US-09-980-370-5 (1-402) x P72196 (1-1097)

```

Accession	Protein Name	Length	Score	E-value	Identity	Positives	Mismatches	Gaps
OY	1 CGAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	522 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATGAGTGAAGTGGCTGCTGCTGCTGCTTCACGACATTTGGAC	1732	97.81%	1.1e-125	97	254	2	0
OY	1 GCAGACTTCACGGAAACCTTCGAGCTTCCTACTCATGAGAGAGACACGACGGAATGAGCT	1732	3.3e-62	1.1e-125	99	254	2	0
OY	1157 ALaasPheRtrIuThpPheGluSeSerThrRtsIcYgIuaIaProIaGluTrpThr	1732	99.25%	1.1e-125	99	254	2	0
OY	61 ACATATCGATGCGGATGCGATGATG							

QY	1177	Thrl1asapalaaspcl	ysapgl	inglytrp	leucyls	leu	ser	scyl	gln	leu	asp	1196
OY	121	TGGCTACAGCTATGCGGCGACCAAGCTACGATCGCTTCCTTCACGATGCAATGCAATGGCC										180
DY	1197	TRPLEUTRRLAHISGLIGYISERSAVALISESERPSESERTRPAISGLIYETALA										1216
OY	181	TTGATCTCGATTAACCTATCTCATCTATCCAAAGCATGTTACAGCGCGCAACAGCAATGAATGAC										240
DY	1217	LEUSTRPACSPASPSMTSLHLLSSGGLYSAPALATHGLGATLTHLYVALILYSTYR										12356
OY	241	TACATCATCGACCAAGAGGGTTTCCGCGGATACATCGATCGCATATGTCATCAACGAC										300
DY	1237	TYRTYALVALASVALASMSAGLYPHEPGTGYASPISTYALVALAMELILISORTYTH										1256
OY	301	GGCGACGACGGCGGACCTTACACGGCTTCCTTTCACAAACGGCTACAGCATATATAC										360
DY	1257	GLYHSAALALACIYASPHETHRYVALVALPHEGLUUTRPIAASGLIYILENLYS										1276
OY	361	GGCGACGACCAATCGCTGCTTTCACAGGACGCAATGGCGCC										402
DY	1277	GLYALVALAAPHEDYLEUSERTHGLVALASNYALTA										1290
RESULT 6												
ID	O51817	PRELIMINARY: PRT: 1732 AA.										
AC	O51817	1996	(TEMBLRef: 01. Created)									
DT	O1-NOV-1996		(1-NOV-1996)									
DT	01-JUN-2002		(TEMBLrel: 21. Last annotation update)									
DT	Porphyralin											
CN	PRP											
OS	Porphyromonas gingivalis (Bacteroides gingivalis)											
OC	Bacteria: CFB group: Bacteroidales: Bacteroidales: Porphyromonadaceae:											
OX	NCBI TaxID=837											
RN	111	SEQUENCE FROM N.A.										
RP	STRAIN=112											
RX	MEDLINE=66213011: PubMed=8631659;											
RT	Proteinase from <i>Porphyromonas gingivalis</i> M8 N., Patel J.M., Whitlock J.,											
RA	Proteinase from <i>Porphyromonas gingivalis</i> M8 N., Patel J.M., Whitlock J.,											
RT	Analysis of the prp gene encoding porphyralin, a cysteine proteinase											
RT	of <i>Porphyromonas gingivalis</i> M8 N., Patel J.M., Whitlock J.,											
RL	J. Bacteriol. 178:2734-2741(1996).											
DR	EMBL: 042210: AA80565.1: -											
DR	Interp: C35: P000977: DNA: 1199se											
DR	Interp: IPH001769: Peptidase_C25.											
DR	Pfam: PF01364: Peptidase_C25: 3											
DR	PROSITE: PS00697: DNA_LGASE_A1: UNKNOWN_1.											
SO	SEQUENCE 1732 AA: 187875 MW: 654271DBEE7BCAE4 CRC64:											
Alignment Scores:												

[illegible]

OY 241 TACTATCCAGTCAGCAAGCGGTTTCCGGGATCACTATGCGGTGATGATTCGAAGAC 300
 Db 1237 TTTTAAAlaValaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1256
 OY 301 GGCAGCAAGCGCGGACGACTTCAAGGTTGTTTCGAAGAAGCCCTTAACGGAATGAATG 360
 Db 1257 GYTThraSnaIaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 1276
 OY 361 GGGCGAGCAAGATTCGCTTTCACAGCAAGCAATGGCGCC 402
 Db 1277 GYgylAlaIaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1290

RESULT 8

O51839 PRELIMINARY; PRT: 1706 AA.
 ID 051839
 AC 051839; 051840: (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE Arginine-specific thiol protease precursor.
 GN Porphyromonas gingivalis (Bacteroides gingivalis).
 OS Bacteria; CFM group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_Taxid=837;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50;
 RX MEDLINE=95160709; PubMed=7857299;
 RA Kirsbaum L., Sotiropoulos C., Jackson C., Cleal S., Stakeski N.,
 RA Reynolds E.C.;
 RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
 RT gingivalis M50 encoding a 132 kDa protein that contains an arginine-
 RT specific thiol endopeptidase domain and a haemagglutinin domain.";
 RL Biochem. Biophys. Res. Commun. 207:424-431(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50;
 RX MEDLINE=96311339; PubMed=8713096;
 RA Stakeski N., Cleal S.M., Reynolds E.C.;
 RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes
 RT an arginine-specific thiol proteinase and multiple adhesins.";
 RL Biochem. Biophys. Res. Commun. 224:603-610(1996).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50;
 RA Reynolds E.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50;
 RA Stakeski N.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L26341; AAC18876.1; -.
 DR HSSP: P95493; ICVR.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Protease; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 228 1706 ARGININE-SPECIFIC THIOI. PROTEASE.
 SO SEQUENCE 1706 AA; 185626 MW; E8BD07C981B844 CRC64;

Alignment Scores:

Pred. No.: 4.13e-62 Length: 1706
 Score: 714.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.67% Indels: 0
 Db: 2 Gaps: 0

US-09-980-370-5 (1-402) x O51839 (1-1706)

OY 1 GCAGACTCCAGCAAGCGGTTTCCGGGATCACTATGAGAGGACCGCAATGGACT 60
 Db 1139 AlaAspPheThrGluThrPheGluSerThrHisGlyGluAlaProAlaGluThrPthr 1158
 OY 61 ACTATGATGCGGATGGGATGAGGTTGGGCTCTGCTCTGCTTCGCAATGGAC 120
 Db 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyThrLeuGlyLeuSerSerGlyGlnLeuAsp 1178
 OY 121 TGGCTCAGAGCTCAGGCGGACCACTAGTAAGCTTCTTCATGGAATGAATGGCT 180
 Db 1179 ThrLeuThrAlaHisGlyGlyThrAsnValAlaSerSerPheSerThrPasnGlyMetAla 1198
 OY 181 TTGAATCCTGATTAATCTATCTCATCTCAAGAGATGTTACAGCGGACCAAGAGTAAGTAC 240
 Db 1199 LeuAsnProAspAspThrLeuIleSerLysAspAlaThrGlyAlaThrLysValLysTyr 1218
 OY 241 TACTATCCAGTCAGCAAGCGGTTTCCGGGATCACTATGCGGATGATTCGAAGAC 300
 Db 1219 TTTTAAAlaValaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1238
 OY 301 GGCAGCAAGCGCGGACGACTTCAAGGTTGTTTCGAAGAAGCCCTTAACGGAATGAATG 360
 Db 1239 GYTThraSnaIaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 1258
 OY 361 GGGCGAGCAAGATTCGCTTTCACAGCAAGCAATGGCGCC 402
 Db 1259 GYgylAlaIaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1272

RESULT 9

P72197 PRELIMINARY; PRT: 1723 AA.
 ID P72197
 AC P72197;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DE Lys-gingipain.
 GN KGP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFM group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_Taxid=837;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
 RA Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
 RT "Molecular cloning and characterization of Porphyromonas gingivalis
 RT Lys-gingipain.";
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U54691; AAA99810.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1723 AA; 186831 MW; 4508A7E50197CEBD CRC64;

Alignment Scores:

Pred. No.: 8.19e-62 Length: 1723
 Score: 711.00 Matches: 130
 Percent Similarity: 99.25% Conservative: 3
 Best Local Similarity: 97.01% Mismatches: 1
 Query Match: 97.26% Indels: 0
 Db: 2 Gaps: 0

US-09-980-370-5 (1-402) x P72197 (1-1723)

OY 1 GCAGACTCCAGCAAGCGGTTTCCGGGATCACTATGAGAGGACCGCAATGGACT 60
 Db 1139 AlaAspPheThrGluThrPheGluSerThrHisGlyGluAlaProAlaGluThrPthr 1158
 OY 61 ACTATGATGCGGATGGGATGAGGTTGGGCTCTGCTCTGCTTCGCAATGGAC 120
 Db 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyThrLeuGlyLeuSerSerGlyGlnLeuAsp 1178
 OY 121 TGGCTCAGAGCTCAGGCGGACCACTAGTAAGCTTCTTCATGGAATGAATGGCT 180
 Db 1179 ThrLeuThrAlaHisGlyGlyThrAsnValAlaSerSerPheSerThrPasnGlyMetAla 1198
 OY 181 TTGAATCCTGATTAATCTATCTCATCTCAAGAGATGTTACAGCGGACCAAGAGTAAGTAC 240
 Db 1199 LeuAsnProAspAspThrLeuIleSerLysAspAlaThrGlyAlaThrLysValLysTyr 1218
 OY 241 TACTATCCAGTCAGCAAGCGGTTTCCGGGATCACTATGCGGATGATTCGAAGAC 300
 Db 1219 TTTTAAAlaValaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1238
 OY 301 GGCAGCAAGCGCGGACGACTTCAAGGTTGTTTCGAAGAAGCCCTTAACGGAATGAATG 360
 Db 1239 GYTThraSnaIaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 1258
 OY 361 GGGCGAGCAAGATTCGCTTTCACAGCAAGCAATGGCGCC 402
 Db 1259 GYgylAlaIaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1272


```

Db      1176 ThrLeaspAlaaspGlyAspGly61nglyTrpleucysLeuSerSerGlyGlnLeuasp 1195
QY      121  TGGCTCAGCTCATGGCGGACCAACGTAAGCTCTTCTCATGCAATGAATGGCT 180
Db      1196 TrpleuthrAlaHisGlyGlyTrhrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1215
QY      181  TTGAATCCGTAACTATCTCATCTCAAGAGATGTTACAGCGGCACGAAAGTAAGTAC 240
Db      1216 LeuAsnProAspAsnTyrlleuileSerLysAspValThrGlyAlaThrLysValLysTyr 1235
QY      241  TACTATCCAGTCACAGCAGGTTTCCCGGGGATCTACTATGCGGTATGATTCACAGAG 300
Db      1236 TyrlTyrlAlaValAlaAsnAspGlyPheProGlyAspHisTyrlAlaValMetLileSerLysThr 1255
QY      301  GGCACGACGACGCGGACTTCACGCTTGTTCGAGAAACCCCTAAGCAATGAATGAAG 360
Db      1256 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyLeuAsnLys 1275
QY      361  GCGCGACGACAGATTCGGTCTTTCACAGCAAGCCCAATGGCGCC 402
Db      1276 GlyGlyAlaArxPheGlyLeuSerThrGluAlaAspGlyAla 1289

```

RESULT 10

```

P96967 PRELIMINARY: PRT: 1358 AA.
AC P96967:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hemagglutinin.
GN HAAGD.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA Han N., Lepine G., Whitlock J., Wojciechowski L., Progulskie-Fox A.;
RT "Cloning, sequencing and characterization of hagd, a member of the
RT HAREP multigene family in Porphyromonas gingivalis.";
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U68468; AAB49691.1;
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1358 AA; 147102 MW; 47FCA0B25B06DEDB8 CRC64;

```

```

Alignment Scores:
Pred. No.: 1 24e-61 Length: 1358
Score: 709.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.99% Indels: 0
DB: 2 Gaps: 0

```

US-09-980-370-5 (1-402) x P96967 (1-1358)

```

QY      1  GCAGACTTCACGAAACGTTGAGTCTTCTACTCATGAGAGGACCAACGCGAATGCACT 60
Db      791  AlaAspPheThrGluThrPheGluSerSerThrHisGlyGluAlaProAlaGluTrpThr 810
QY      61  ACTATGATCGCGATGGCGATGGTGGAGGGTGGCTGTCTGCTCCGCAATGGAC 120
Db      811  ThrLeaspAlaAspGlyAspGlyGlnAspTrpleucysLeuSerSerGlyGlnLeuasp 830
QY      121  TGGCTCAGCTCATGGCGGACCAACGTAAGCTCTTCTCATGCAATGAATGGCT 180
Db      831  TrpleuthrAlaHisGlyGlyTrhrAsnValValAlaSerPheSerTrpAsnGlyMetAla 850
QY      181  TTGAATCCGTAACTATCTCATCTCAAGAGATGTTACAGCGGCACGAAAGTAAGTAC 240

```

```

Db      851  LeuAsnProAspAsnTyrlleuileSerLysAspValThrGlyAlaThrLysValLysTyr 870
QY      241  TACTATCCAGTCACAGCAGGTTTCCCGGGGATCTACTATGCGGTATGATTCACAGACG 300
Db      871  TyrlTyrlAlaValAlaAsnAspGlyPheProGlyAspHisTyrlAlaValMetLileSerLysThr 890
QY      301  GGCACGACGACGCGGACTTCACGCTTGTTCGAGAAACCCCTAAGCAATGAATGAAG 360
Db      891  GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyLeuAsnLys 910
QY      361  GCGCGACGACAGATTCGGTCTTTCACAGCAAGCCCAATGGCGCC 402
Db      911  GlyGlyAlaArxPheGlyLeuSerThrGluAlaAspGlyAla 924

```

RESULT 11

```

P72194 PRELIMINARY: PRT: 1723 AA.
AC P72194:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT "Cloning and sequencing of the gene encoding a novel lysine-specific
RT cysteine proteinase (Lys-gingipain) in Porphyromonas gingivalis:
RT structural relationship with the arginine-specific cysteine proteinase
RT (Arg-gingipain).";
RL J. Biochem. 120:398-406(1996).
DR EMBL: D83258; BAA11870.1;
DR MEROPS: C25.002;
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1723 AA; 187261 MW; 5628963D251493EB CRC64;

```

```

Alignment Scores:
Pred. No.: 1 29e-61 Length: 1723
Score: 709.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.99% Indels: 0
DB: 2 Gaps: 0

```

US-09-980-370-5 (1-402) x P72194 (1-1723)

```

QY      1  GCAGACTTCACGAAACGTTGAGTCTTCTACTCATGAGAGGACCAACGCGAATGCACT 60
Db      1156  AlaAspPheThrGluThrPheGluSerSerThrHisGlyGluAlaProAlaGluTrpThr 1175
QY      61  ACTATGATCGCGATGGCGATGGTGGAGGGTGGCTGTCTGCTCCGCAATGGAC 120
Db      1176  ThrLeaspAlaAspGlyAspGlyGlnAspTrpleucysLeuSerSerGlyGlnLeuasp 1195
QY      121  TGGCTCAGCTCATGGCGGACCAACGTAAGCTCTTCTCATGCAATGAATGGCT 180
Db      1196  TrpleuthrAlaHisGlyGlyTrhrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1215
QY      181  TTGAATCCGTAACTATCTCATCTCAAGAGATGTTACAGCGGCACGAAAGTAAGTAC 240
Db      1216  LeuAsnProAspAsnTyrlleuileSerLysAspValThrGlyAlaThrLysValLysTyr 1235
QY      241  TACTATCCAGTCACAGCAGGTTTCCCGGGGATCTACTATGCGGTATGATTCACAGACG 300

```


Qy	1236	TytrYrLrAlaValaInsaSpGlyPheProGlyAspHisTyrLrAlaValMeLIEseLysThr	1255
Db	1256	GlyThrAsnAlaGlyAspPheThrValaIAlaPheGluGluThrProAsnGlyIleAsnLys	1275
Oy	361	GGCGGAGCAAGATTGGCTCTTTCACGGAGGCCAATGGCGCC	402
Db	1276	GlyGlyAlaArGpPheGlyLeuSerThrGluAlaInslGlyAla	1289
RESULT 12			
ID	051838	PRELIMINARY:	PRT: 1706 AA.
AC	051838:		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Protease precursor.		
CN	PRPRL		
OS	Porphyromonas gingivalis (Bacteroides gingivalis).		
OC	Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;		
OC	Porphyromonas.		
OX	NCBI_TaxID=837;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W50;		
RC	MEDLINE=96071894; PubMed=7591131;		
RA	Aduse-Opoku J., Muir J., Staney J.M., Rangarajan M., Curtis M.A.;		
RT	"Characterization, genetic analysis, and expression of a protease		
RT	antigen (PrpRL) of Porphyromonas gingivalis W50."		
RL	Infect. Immun. 63:4744-4754(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W50;		
RA	Rangarajan M., Aduse-Opoku J., Staney J.M., Young K.A., Curtis M.A.;		
RT	"The prpRL and the prp2 arginine-specific protease genes of		
RT	Porphyromonas gingivalis W50 produce five biochemically distinct		
RT	enzymes."		
RL	Mol. Microbiol. 23:0-0(1997).		
DR	EMBL: X82680; CA57997.1; -.		
DR	HSSP: P95493; ICVR.		
DR	MEROPS: C25.001; -.		
DR	InterPro: IPR000977; DNM_11gase.		
DR	InterPro: IPR001769; Peptidase_C25.		
DR	Pfam: PF01364; Peptidase_C25; 3.		
DR	PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.		
KW	Signal.		
FT	SIGNAL. 1 23	POTENTIAL.	
FT	CHAIN 228 719	ALPHA-PROTEASE.	
FT	CHAIN 720 1262	BETA-ADHESIN.	
SO	SEQUENCE 1706 AA; 185705 MW; 0E56DCD87FDA8CDD CRC64;		
Alignment Scores:			
Pred. No.:	2.55e-61	Length:	1706
Score:	706.00	Matches:	130
Best Similarity:	98.51%	Conservative:	2
Percent Similarity:	97.01%	Mismatches:	2
Query Match:	96.58%	Indels:	0
DB:	2	Gaps:	0
US-09-980-370-5 (1-402) x 051838 (1-1706)			
Oy	1	GCAGACTTCACGGAACGTTGAGCTCTACTCATGAGAGGACACGCGAATGACT	60
Db	1139	AlaAspPheThrGluThrPheGluSerSerThrHisGlyAlaIleThrAlaGluTrpThr	1158
Oy	61	ACTATGATGCCGATGCCGATGCTGAGGGTTCGCTCTGCTTCGGAATTGAC	120
Db	1159	ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerGlyGlnLeuAsp	1178
Oy	121	TGGCTCACAGCTCATGGCGGACACCAAGTAGTAAGCTTTCTCATGGAATGGATGCT	180

Dd	1179	TripleutheralahisglgylYThrnsnValValSerSerPheSerTrpAsnGlyMetAla	1199
Qy	181	TTGAATCTCTGATACATATCTATCTATCTCAAGAGTGTTCACAGGCGACCAAGAGTAAAGTAC	240
Dd	1199	LeuasnProbaAspAsnTyrLeuIleSerLysAspValTrnGlyAlaThrLysValLysTyr	1210
Qy	241	TACATATCCAGTCACAGCAGCGGTTTTCCCGGGGTCACCTTACGGCGATATCTCCAAAGAGC	300
Dd	1219	TyTyrAlaValaValasnAspGlyPheProGlyasnPhsIstrYAlaValaIleIleSerLysThr	1233
Qy	301	GGCAGCAAGCGCGGAGACTTCACGGCTTTTTCGAAAGAACGCCCTACAGGAAATAAATAG	360
Dd	1239	GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyLeuasnLys	1255
Qy	361	GGCGGAGCAACATTCGCTGCTTTCCACGCAAGCCAAATGGCGCC	402
Dd	1259	GlyGlyAlaArgPheGlyLeuSerThrGluAlaAspGlyAla	1272

RESULT 13

Q9F4J0	PRELIMINARY:	PRT:	925	AA.
AC	Q9F4J0:			
Dt	01-MAR-2001 (TREMBLrel. 16, Created)			
Dt	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
Dt	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Putative outer membrane protein p657.			
OS	Porphyromonas gingivalis (Bacteroides gingivalis).			
OC	Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;			
OX	Porphyromonas.			
OX	NCBI_TaxID=837;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=J50:			
RA	Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,			
RA	Hocking D., Webb E.;			
RT	"P. gingivalis polypeptides and nucleic acids."			
RL	Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.			
RT	Porphyromonas gingivalis."			
RL	Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AY007522; AAG24228.1; "			
DR	InterPro: IPR003961; FN_III.			
DR	SMART: SM00060; FN3; 3			
SD	SEQUENCE 925 AA: 103632 MW: 55F2198D691ADAE8 CRC64:			

Alignment Scores:

Pred. No.:	8,25e-12	Length:	925
Score:	205.00	Matches:	55
Percent Similarity:	47.26%	Conservative:	14
Best Local Similarity:	37.67%	Mismatches:	43
Query Match:	28.04%	Indels:	34
DB:	2	Gaps:	7

US-09-980-370-5 (1-402) x Q9F4J0 (1-925)

Qy	1	GCAGACTTCACGGAACGTTGAGTCTTCTACT	33
Dd	288	AlaValTyrAspGluSerIleGluSerSerThrValCysGlyYThrLeuHisTyrAlaIthr	307
Qy	34	-----CATGGAGAGGCAACGCGGAATGGACTACTATC	66
Dd	308	AspAlaIleLeuTyrGluAsnDheGluasnGlyProValProAsnGlyTyrIleuValIle	327
Qy	67	CATGCCGATGCGATGCTAGAGGTTGGACTCTGCTGTCTTCCGACAA	114
Dd	328	AspAlaAspGlyAspGlyPheSerTrp-----GlyHisTyrLeuAsnAla	342
Qy	115	TTTGACCTGGCTCAGACCTCATGGCGGCAACAGTAGTAAGCTTTTCTCATGCG	171

```

Db      343 TYRAspAlaPheProGlyHisAsnGlyGlyHisCysSerLeuSerAlaSerTyrValPro 362
QY      172 GGAAATGGCT---TTGAAATCCTGTAATATCTATCTCAAGAGATGTTACAGCGCGCAACG 228
Db      363 GlyIleGlyProValThrProAspAsnTyrLeuIleThrProLysValGluGlyAlaLys 382
QY      229 AAGGTAAATGTAATCACTATCCAGTCAACGAC---GGTTTCCCGGCGATCCATATGCGGTG 285
Db      383 ArgValLysTyrTrpValSerThrGlnAspAlaAsnTrpAlaLysGluHisTyrAlaVal 402
QY      286 ATGATCTCCAGACGCGCACGCAAGCGGAGACTTCAGGTTTTCGAAAGAACG--- 342
Db      403 MetAlaSerThrThrGlyThrAlaValGlyAspPheValIleLeuPheGluGluThrMet 422
QY      343 -----CGTAAGCA 351
Db      423 ThrAlaLysProThrGly 428

```

RESULT 14

```

O9KIB3 PRELIMINARY: PRT: 312 AA.
AC 09KIB3:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical outer membrane protein PG27.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CF8 group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas
NCBI_TaxID=837;

```

```

RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=MS0;
RA Rose B., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,
RA Hocking D., Webb E.,
RT "P. gingivalis polypeptides and nucleic acids.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
(2)
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=MS0;
RA Rose B.C.;
RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF237555; AAF81413.1;
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 2.
SQ SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25F5 CRC64;

```

```

Alignment Scores:
Pred. No.: 2,07e-09 Length: 312
Score: 180.00 Matches: 50
Percent Similarity: 55.46% Conservative: 16
Best Local Similarity: 42.02% Mismatches: 35
Query Match: 24.62% Indels: 18
Gaps: 7

```

US-09-980-370-5 (1-402) x O9KIB3 (1-312)

```

QY      10 ACGGAAGCTTGGAGCTTCTTCTATGAGAGAGACGACGGAATGAGTACTATGAT 69
Db      5 SerGluSerPheGluSerGly-----IleProAlaIleTrpLysThrIleAsp 20
QY      70 GCCCATGGCGATGAGTGGAGGCTTGGCTGTCTGTCTCCGACAATTTGAGTGGCTACCA 129
Db      21 AlaAspGlyAspGlyTyrAsnTrpMetHisLeuThrAsn-----PheThr 35
QY      130 GCTATGCGCGACCAACAGTAAAGCTTTTCATGGAATGA-----ATGCGCTTG 183
Db      36 GlyIleGlyLeu---CysValSerSerAlaSerTyrIleGlyValGlyAlaLeu 54
QY      184 AATCGTGAATCAATTCATCTCAAGAGATGTTACAGCGCGCAACG-----AAG 231
Db      184 AATCGTGAATCAATTCATCTCAAGAGATGTTACAGCGCGCAACG-----AAG 231

```

```

Db      55 ThrProAspAsnTyrIleuIleThrProGluLeuLysLeuProThrAspAlaLeuValGlu 74
QY      232 GTAAGTACTACTATTCACAGTCAACGAC---GGTTTCCCGGCGATCACTATGCGGTATG 288
Db      75 IleIleTyrTrpValCysThrGlnAspLeuThrAlaProSerGluHisTyrAlaValTyr 94
QY      289 ATCTCCAGACGGGCGACGACGACCGGAGACTTC---ACGTTTTCGAAAGAACG 342
Db      95 SerSerSerThrGlyAsnAsnAlaAlaAspPheValAsnLeuLysGluGluThr 113

```

RESULT 15

```

O9JZ15 PRELIMINARY: PRT: 218 AA.
ID O9JZ15:
AC O9JZ15:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Isomerase, putative.
GN NMB1338.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
NCBI_TaxID=491;

```

```

RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Ulterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.W., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
DR EMBL: AE002482; AAF41713.1;
DR TIGR: NMB1338;
DR InterPro: IPR002529; FAA_hydrolase.
DR Pfam: PF01557; FAA_hydrolase; 1.
KW Complete proteome.
SQ SEQUENCE 218 AA; 23058 MW; BA531FCD9C010BD9 CRC64;

```

```

Alignment Scores:
Pred. No.: 1.39 Length: 218
Score: 90.50 Matches: 38
Percent Similarity: 36.36% Conservative: 14
Best Local Similarity: 26.57% Mismatches: 48
Query Match: 12.38% Indels: 43
Gaps: 6

```

US-09-980-370-5 (1-402) x O9JZ15 (1-218)

```

QY      67 GATCCGATGGCGATGAGGAT----- 90
Db      77 AspSerAspGlyThrGlnGluGlyLysAspIleLeuGlyCysValAlaGlyTyrGlyVal 96
QY      91 -----TGCGTCTGTCTGTCTTCCGGACAATTTGAGCTGG 123
Db      97 GlyLeuAspLeuThrAlaArgAspIleGlnCysArgLeuLysGlyLeuProThr 116
QY      124 CTCACAGCTCATGGC-----GGACCAACAGTAAAGCTTTTCATGGAATGGAATG 177
Db      117 LeuLysAlaLysGlyPheArgHisSerAlaCysValSerAspPheAlaAlaGlyArg 136
QY      178 GCTTGAATCCTGATTAATATATCATGCAAGAGATGTTACAGCGCGCAAGGTAAG 237
Db      137 IleGlyAsnProGluLysValLeuPheSerLeuLysGlnAsnGlyValLeuLysGlnArg 156
QY      238 -----TACTATATTCACAGTCAACGAGGTTTCCGGGAGATCACTATGCG 282
Db      157 GlyAspThrGlyLeuMetIleTyrProIleArgGlu----- 168

```

```

OY      283 GTGATGATCTCCAGACG-----GCCACGAACGCCGAGACTTCACGGTTGTT 330
          ::::: |||
Db      169 ---IleLeuH1sLysLeuAlaAlaAspTyrGlyLeuGlyLysGlyAsp-----LeuVal 185
          ::::: |||
OY      331 TTGGAAGAAACGCCCTTACGGAATTAATTAAGGCGGACGACGATTCGGTCTTTCCACGGAA 390
          ||| ||| ::::: |||
Db      186 PheThrGlyThrProSerGlyValGlyAlaIleGlyAlaGlyAspAsnLeuAlaLeuGlu 205
          ::::: |||
OY      391 GCCAATGGC 399
          ::::: |||
Db      206 LeuAspGly 208

```

Search completed: June 26, 2003, 00:44:14
 Job time : 44.5 secs

!

Thu Jun 26 11:58:13 2003

us-09-980-370-6.frag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:59 : Search time 47.2941 Seconds
(without alignments)
377.543 Million cell updates/sec

Title: US-09-980-370-6
Perfect score: 730
Sequence: 1 ADPTEPSTHCEAPAEWT.....PNGINKGARELSTENGCA 134

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database:

1: A.Geneseq.101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	730	100.0	134	22	AA849217
2	716	98.1	1687	17	AA86033
3	716	98.1	1687	19	AA86945
4	716	98.1	1704	16	AA87048
5	716	98.1	1704	18	AA87183
6	716	98.1	1704	21	AA88336
7	716	98.1	1704	22	AA88603
8	716	98.1	1704	22	AA88603
9	716	98.1	1704	22	AA88603
10	716	98.1	1704	22	AA88603

11	714	97.8	1706	18	AA824786	PRR antigenic pro
12	711	97.4	439	17	AA86024	P. gingivalis haem
13	711	97.4	439	19	AA86942	Haemagglutinin pro
14	711	97.4	450	17	AA86021	P. gingivalis haem
15	711	97.4	450	19	AA86948	Haemagglutinin pro
16	711	97.4	450	19	AA86948	Haemagglutinin pro
17	711	97.4	450	19	AA86948	Haemagglutinin pro
18	709	97.1	1087	17	AA86948	Haemagglutinin pro
19	709	97.1	1087	17	AA86948	Haemagglutinin pro
20	709	97.1	1087	17	AA86948	Haemagglutinin pro
21	709	97.1	1087	17	AA86948	Haemagglutinin pro
22	704	96.4	436	17	AA86948	Haemagglutinin pro
23	704	96.4	436	17	AA86948	Haemagglutinin pro
24	704	96.4	436	17	AA86948	Haemagglutinin pro
25	704	96.4	436	17	AA86948	Haemagglutinin pro
26	704	96.4	436	17	AA86948	Haemagglutinin pro
27	704	96.4	436	17	AA86948	Haemagglutinin pro
28	704	96.4	436	17	AA86948	Haemagglutinin pro
29	704	96.4	436	17	AA86948	Haemagglutinin pro
30	704	96.4	436	17	AA86948	Haemagglutinin pro
31	704	96.4	436	17	AA86948	Haemagglutinin pro
32	704	96.4	436	17	AA86948	Haemagglutinin pro
33	704	96.4	436	17	AA86948	Haemagglutinin pro
34	704	96.4	436	17	AA86948	Haemagglutinin pro
35	704	96.4	436	17	AA86948	Haemagglutinin pro
36	704	96.4	436	17	AA86948	Haemagglutinin pro
37	704	96.4	436	17	AA86948	Haemagglutinin pro
38	704	96.4	436	17	AA86948	Haemagglutinin pro
39	704	96.4	436	17	AA86948	Haemagglutinin pro
40	704	96.4	436	17	AA86948	Haemagglutinin pro
41	704	96.4	436	17	AA86948	Haemagglutinin pro
42	704	96.4	436	17	AA86948	Haemagglutinin pro
43	704	96.4	436	17	AA86948	Haemagglutinin pro
44	704	96.4	436	17	AA86948	Haemagglutinin pro
45	704	96.4	436	17	AA86948	Haemagglutinin pro

ALIGNMENTS

RESULT 1	AA849217	standard; protein: 134 AA.
ID	AA849217	
AC	AA849217	
DT	13-MAR-2001	(first entry)
DE	Peptide used in the invention.	
DE	HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral.	
KW	Unidentified.	
OS	Unidentified.	
PN	WO200072875-A1.	
XX	07-DEC-2000.	
XX	26-MAY-2000: 2000WO-AU00599.	
XX	28-MAY-1999: 99AU-0000652.	
XX	(UNSY) UNIV SYDNEY.	
XX	Colliger CA, Hunter N, De Carlo MA.	
XX	WPI: 2001-080424/09.	
XX	Treating microbial infection in environment containing porphyrin, by	
XX	administering a HA-2 antagonist	
XX	Claim 9; Page 98-99; 102pp; English.	

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the iron, heme
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an H42 domain, and an H42-binding
 CC moiety on a porphyrin containing molecule present in the environment.
 CC The present invention also relates to a method for the prophylaxis and
 CC treatment of a disease, such as periodontal disease, by the use of a
 CC vaccine against a microorganism, such as Porphyromonas gingivalis, resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

SO Sequence 134 AA:

Query Match 100.0%; Score 730; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 7.9e-68;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADFTFFESSHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGMA 60
 DB 1 ADFTFFESSHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGMA 60
 OY 61 LNPDNLTLSKDVGTATKVKYYPVNDGFGPDHYAVMISKTGTAGDFTVVEETPNCINK 120
 DB 61 LNPDNLTLSKDVGTATKVKYYPVNDGFGPDHYAVMISKTGTAGDFTVVEETPNCINK 120
 OY 121 GCARFGLSTEANGA 134
 DB 121 GCARFGLSTEANGA 134

RESULT 2
 AAR96033 standard; Protein: 1687 AA.

XX ID AAR96033:
 AC AAR96033:
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haemagglutinin hage.
 KW Haemagglutinin; hage; periodontal disease; vaccine; antibody.
 OS Porphyromonas gingivalis strain FDC381.
 PN WC9617936-A2.
 PD 13-JUN-1996.
 PR 11-DEC-1995; 95WO-US16108.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepline G, Patel JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI NPI: 1996-287181/29.
 DR N-PSDB: AAT30656.
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 CC and vaccination against periodontal disease
 CC
 CC Claim 5: Page 138-143; 153pp; English.
 CC P. gingivalis 381 haemagglutinin hage (AAR96033) was identified as
 CC the product of a gene (AAT30656) identified in P. gingivalis 318
 CC genomic DNA. The haemagglutinin can be obt'd. from transformed
 CC host cells and used as a vaccine to protect humans or animals against
 CC periodontal disease. Expression in Salmonella cells allows prodn. of
 CC a live vaccine. The haemagglutinin can also be used to detect the

CC presence of anti-P. gingivalis antibodies and to raise monoclonal
 CC antibodies for diagnostic appln.

SO Sequence 1687 AA:

Query Match 98.1%; Score 716; DB 17; Length 1687;
 Best Local Similarity 97.8%; Pred. No. 7.1e-63;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTFFESSHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGMA 60
 DB 1120 ADFTFFESSHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGMA 1179
 OY 61 LNPDNLTLSKDVGTATKVKYYPVNDGFGPDHYAVMISKTGTAGDFTVVEETPNCINK 120
 DB 1180 LNPDNLTLSKDVGTATKVKYYPVNDGFGPDHYAVMISKTGTAGDFTVVEETPNCINK 1239
 OY 121 GCARFGLSTEANGA 134
 DB 1240 GCARFGLSTEANGA 1253

RESULT 3
 AAM69495 standard; Protein: 1687 AA.

XX ID AAM69495:
 AC AAM69495:
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein hage.
 KW Haemagglutinin protein; periodontal disease; vaccine; hage.
 OS Porphyromonas gingivalis.
 PN US5824791-A.
 PD 20-OCT-1998.
 PR 11-DEC-1995; 95US-0570311.
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepline G, Patel JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI N-PSDB: AAV58881.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 CC and/or protease poly:peptide(s)
 CC
 CC Claim 1: Column 167-182; 101pp; English.
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hage haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

SO Sequence 1687 AA:

Query Match 98.1%; Score 716; DB 19; Length 1687;
 Best Local Similarity 97.8%; Pred. No. 7.1e-63;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSNMGMA 60
Db 1120 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSNMGMA 1179
QY 61 LNPDNVILISKDVATGATVKYKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 120
Db 1180 LNPDNVILISKDVATGATVKYKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 1239
QY 121 GGARFGLSTEANGA 134
Db 1240 GGARFGLSTEANGA 1253

RESULT 4
AAR70188

ID AAR70188 standard; protein; 1704 AA.

AC AAR70188;

DT 21-SEP-1995 (first entry)

DE Arg-gingipain-2 prepolyprotein.

KW Arg-gingipain-2; gingivitis; periodontal disease; vaccine;
KM arginine-specific protease.

XX Porphyromonas gingivalis.

FH Key Location/Qualifiers

FT 228..719 /label= Protease

FT /note= "corresponds to Arg-gingipain-1"

FT 720..1091

FT Region /label= Hemagglutinin

FT 1092..1429 /label= Hemagglutinin

FT 1430..1704 /label= Hemagglutinin

FT Region

XX MO9507286-A.

XX 16-MAR-1995.

XX 09-SEP-1994; 94MO-US10283.

XX 24-JUN-1994; 94US-026541.

XX 10-SEP-1993; 93US-0119361.

XX 21-OCT-1993; 93US-0141324.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Barr PJ, Pavloff N, Potempa J, Travis J;

XX MPI; 1995-12373/16.

XX N-PSDB; AAO83489.

XX DNA encoding Arg-gingipain proteins - used to develop prods. for

XX detection, treatment and prevention of periodontal disease

XX Disclosure; Page 70-77; 89pp; English.

XX A low mol. wt. arginine-specific gingipain (AG-1) and high mol. wt.

XX AG (AG-2) were isolated from P. gingivalis strains H66 (ATCC

XX 3327) and M50 (ATCC 53973). The sequences of the proteins were

XX used to design PCR primers and probes to isolate AG DNA. Lambda

XX DASH and lambda ZAP libraries were screened with a probe based on

XX amino acids 11-22 of the AG protein to obtain DNA encoding AG-1

XX (AAO83484) and AG-2 (AAO83489). AG-2 is a prepolyprotein

XX incorporated AG-1.

XX Sequence 1704 AA;

Query Match 98.1%; Score 716; DB 16; Length 1704;
Best Local Similarity 97.8%; Pred. No. 7,2e-65;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSNMGMA 60
Db 1137 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSNMGMA 1196
QY 61 LNPDNVILISKDVATGATVKYKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 120
Db 1197 LNPDNVILISKDVATGATVKYKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 1256
QY 121 GGARFGLSTEANGA 134
Db 1257 GGARFGLSTEANGA 1270

RESULT 5
AAW34843

ID AAW34843 standard; protein; 1704 AA.

AC AAW34843;

DT 03-JUN-1998 (first entry)

DE Arg-gingipain high molecular weight prepolyprotein sequence.

KW Arg-specific gingipain protease; gingivitis; periodontal disease;

KM vaccine; infection.

XX Porphyromonas gingivalis.

FH Key Location/Qualifiers

FT 1..227 /note= "precursor protein"

FT MO934629-A1.

XX 25-SEP-1997.

XX 21-MAR-1997; 97WO-US04635.

XX 22-MAR-1996; 96US-0013945.

XX (MORE-) MOREHOUSE SCHOOL MEDICINE.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Genco CA, Potempa J, Travis J, Genco C;

XX MPI; 1997-479993/44.

XX N-PSDB; NAT93872.

XX Porphyromonas gingivalis Arg-specific gingipain protease peptide(s)

XX - useful for protecting animals and humans from gingivitis and

XX periodontal diseases

XX Disclosure; Pages 68-73; 95pp; English.

XX The present sequence represents an arginine-specific protease of

XX Porphyromonas gingivalis. The following peptides, derived from Arg-

XX and Lys-specific high molecular weight proteases, offer protection

XX against infection: YTVVYRDGK IREGLATAT DGVATGNEH YCIEKYTAGS VSPKYC

XX (I): YTPVEEKONG RMIVIVAKKY (II): QLPETPVAC VNGDFEFSMP CEALMNRQ

XX (III): GEPNRPQPS NUTATGCGK VTLKMDAPSK (IV): GNHEVCVEYK VTAGVSPKYC

XX KDVTV (V): RRENNVEPR YTPVEEKONG (VI): TFAEFEDTYR RMPMNEPR (VII):

XX KIEGLTATF EEDG (IX): RDGTIKIEGL TATFEEDGV ATGN (X): KIEGLTATF

XX FEEDGATGN HEY (XI): KMDAPNGPN PNP PNP PNP PNP PNP PNP PNP PNP PNP

XX YTPVEEKONG RMIVIVAKKY (XII). They are used in vaccines to protect

XX animals, including humans, from gingivitis and/or periodontal

XX diseases.

XX Sequence 1704 AA;

Query Match 98.1%; Score 716; DB 18; Length 1704;
 Best Local Similarity 97.8%; Pred. No. 7.2e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 60
 |||
 DB 1137 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 1196
 |||

QY 61 LMPDNVLSKDVYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNCINK 120
 |||
 DB 1197 LMPDNVLSKDVYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNCINK 1256
 |||

QY 121 GGARFGLSTEANGA 134
 |||
 DB 1257 GGARFGLSTEANGA 1270
 |||

RESULT 6
 AAY67396
 ID AAY67396 standard; Protein; 1704 AA.
 XX
 AC AAY67396;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Arg-gingipain-2 amino acid sequence.
 XX
 KW Arginine specific proteinase; Arg-gingipain; gingipain-2; haemagglutinin;
 XX immunogenic component; vaccine; inflammatory response; tissue damage;
 XX periodontal disease.
 XX
 OS Porphyromonas gingivalis.
 XX
 FH Key Location/Qualifiers
 FT Region 229..719
 FT /note= "Amino acids 229-719 are specifically claimed"
 FT Region 720..1185
 FT /note= "Amino acids 720-1185 are specifically claimed"
 XX
 PN US6017532-A.
 XX
 PD 25-JAN-2000.
 XX
 PF 08-NOV-1994; 94US-0336308.
 XX
 PR 10-SEP-1993; 93US-0119361.
 PR 24-JUN-1994; 94US-0265441.
 XX
 PA (UYGE-) UNIV GEORGIA RES FOUND INC.
 XX
 PI Potempa JS, Travis J;
 XX
 DR WPI: 2000-136659/12.
 DR N-PSDB: AAZ60181.
 XX
 PT New Porphyromonas gingivalis arginine-specific protease preparation
 PT useful for preparing vaccines against periodontal disease and for
 PT screening for Arg-gingipain inhibitors
 XX
 PS Claim 1: Column 29-42; 55pp; English.
 XX
 CC This sequence represents a Porphyromonas gingivalis arginine-specific
 CC proteinase known as Arg-gingipain/gingipain-2 amino acid sequence.
 CC Gingipain-2 consists of a 50kD protease component non-covalently
 CC associated with a 44kD haemagglutinin component. The proteinase is
 CC stimulated by glycine containing peptides and glycine analogues. It is
 CC inhibited by cysteine protease group specific inhibitors. The protease
 CC preparation can be used in immunogenic compositions and vaccines against
 CC inflammatory response and tissue damage caused by P. gingivalis in
 CC periodontal disease. It can also be used to screen for agents that
 CC modulate Arg-gingipain proteinase activity inhibitors.

SQ Sequence 1704 AA;
 Query Match 98.1%; Score 716; DB 21; Length 1704;
 Best Local Similarity 97.8%; Pred. No. 7.2e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 60
 |||
 DB 1137 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 1196
 |||

QY 61 LMPDNVLSKDVYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNCINK 120
 |||
 DB 1197 LMPDNVLSKDVYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNCINK 1256
 |||

QY 121 GGARFGLSTEANGA 134
 |||
 DB 1257 GGARFGLSTEANGA 1270
 |||

RESULT 7
 AAU08938
 ID AAU08938 standard; Protein; 1704 AA.
 XX
 AC AAU08938;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE P. gingivalis high molecular weight Arg-gingipain-2.
 XX
 KW Periodontitis; antiinflammatory; Arg-gingipain-2; AG-2; Immunogen.
 XX
 OS Porphyromonas gingivalis.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..227
 FT /label= Prepro-peptide
 FT Region 670..674
 FT /label= Proteolytic-component
 FT Cleavage-site 719
 FT Protein 720..1091
 FT /label= HGP-44kDa
 FT /note= "Haemagglutinin protein component"
 FT Region 599..619
 FT /note= "Region of homology with cysteine proteases"
 FT Cleavage-site 1091
 FT Protein 1092..1429
 FT /label= HGP-17kDa
 FT /note= "Haemagglutinin protein component"
 FT Cleavage-site 1429
 FT Protein 1430..1704
 FT /label= HGP-27kDa
 FT /note= "Haemagglutinin protein component"
 XX
 PN US6274718-B1.
 XX
 PD 14-AUG-2001.
 XX
 PF 25-JAN-2000; 2000US-0490931.
 XX
 PR 24-JUN-1994; 94US-0265441.
 PR 08-NOV-1994; 94US-0336308.
 PR 10-SEP-1993; 93US-0119361.
 PR 09-SEP-1994; 94WO-US10283.
 XX
 PA (UYGE-) UNIV GEORGIA RES FOUND INC.
 XX
 PI Travis J, Potempa JS, Barr PJ, Pavloff N;
 XX
 DR WPI: 2001-588904/66.
 DR N-PSDB: AAS15242.
 XX
 PT New recombinant DNA molecule which encodes high molecular weight
 PT (mature) Arg-gingipain protein, useful for immunisation against

PT Inflammation and tissue damage, comprises enzymatically active protease
 PT component and haemagglutinin component -
 PS Claim 1: Column 29-41: 56pp: English.

XX The invention relates to a recombinant DNA molecule encoding high
 CC molecular weight (mature) Arg-gingipain (AG) protein, which has an
 CC enzymatically active protease component (AG-2) and a haemagglutinin
 CC component, from P. gingivalis. The nucleic acid is useful for producing
 CC mature Arg-gingipain protein. Immunogenic compositions comprising
 CC Arg-gingipain are useful for immunising animals including humans against
 CC inflammatory response and tissue damage caused by an archaeobacterium
 CC Porphyromonas gingivalis, which causes progressive periodontitis.
 CC Arg-gingipain is also useful for identifying agents that modulate
 CC itself or preventing the interaction of the proteinase with the protein
 CC in the gingival area, such as complement factors C3 or C5. The
 CC present sequence is Arg-gingipain-2.

XX Sequence 1704 AA:

Query Match 98.1%; Score 716; DB 22; Length 1704;
 Best Local Similarity 97.8%; Pred. No. 7.2e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTTFESTHGEAPAEWTTIDADGGGWLCLSSGOLDWLTAGGTNNVSSFSNMGA 60

DB 1137 ADFTTFESTHGEAPAEWTTIDADGGGWLCLSSGOLDWLTAGGTNNVSSFSNMGA 1196

OY 61 LNPDNYLISKDVYGAATKRVYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEETPNCINK 120

DB 1197 LNPDNYLISKDVYGAATKRVYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEETPNCINK 1256

OY 121 GGARFGLSTEANCA 134

DB 1257 GGARFGLSTEANCA 1270

RESULT 8
 AAR96029 standard; Protein: 1732 AA.

XX AAR96029;

DT 04-SEP-1996 (first entry)

DE P. gingivalis porphyrein.

KW Porphyrein: haemagglutinin; periodontal disease; vaccine; antibody.

XX Porphyromonas gingivalis strain W12.

OS Location/Qualifiers

FT Region /note="Pro-Asn repeat region type 1"

FT Region /note="Pro-Asn repeat region type 2"

FT Region /note="Pro-Asn repeat region type 1"

FT Region /note="Pro-Asn repeat region type 3"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 2"

FT Region /note="Pro-Asn repeat region type 3"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 2"

PN M09617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995; 95MO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepine G, Patil JM, Progulake-Fox A;

XX Tumasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB; AAT30653.

XX Claim 5: Page 76-81: 153pp: English.

XX P. gingivalis W12 cysteine protease, porphyrein (AAR96029), was
 CC identified as the product of the prt gene (AAT30653) isolated from
 CC P. gingivalis W12 genomic DNA. The porphyrein shows homology to
 CC the haemagglutinins (see also AAR96028-28 and AAR96030-33) of P.
 CC gingivalis 318. It can be obtained from transformed host cells and
 CC used as a vaccine to protect humans or animals against periodontal
 CC disease. Expression in *Salmonella* cells allows production of a live
 CC vaccine. The porphyrein and haemagglutinins can also be used to
 CC detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic applications.

SO Sequence 1732 AA:

Query Match 97.9%; Score 715; DB 17; Length 1732;
 Best Local Similarity 97.8%; Pred. No. 9.4e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTTFESTHGEAPAEWTTIDADGGGWLCLSSGOLDWLTAGGTNNVSSFSNMGA 60

DB 1157 ADFTTFESTHGEAPAEWTTIDADGGGWLCLSSGOLDWLTAGGTNNVSSFSNMGA 1216

OY 61 LNPDNYLISKDVYGAATKRVYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEETPNCINK 120

DB 1217 LNPDNYLISKDVYGAATKRVYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEETPNCINK 1276

OY 121 GGARFGLSTEANCA 134

DB 1277 GGARFGLSTEANCA 1290

RESULT 9
 AAM24787 standard; Protein: 1732 AA.

XX AAM24787;

DT 25-NOV-1997 (first entry)

DE PTK antigenic protein complex.

KW Periodontal disease; cell surface protein; chitin protease;

KW haemagglutinin; adhesin; therapy; diagnosis; vaccine; antigen.

OS Porphyromonas gingivalis strain W50.

FT Region /note="Pro-Asn repeat region type 3"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 2"

FT Region /note="Pro-Asn repeat region type 3"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 2"

FT Region /note="Pro-Asn repeat region type 3"

DE Prrr antigenic protein complex.

XX Periodontal disease: cell surface protein: thiol protease:

KM endopeptidase: Prrr: Prrr45; Prrr15; Prrr17; Prrr27.

KM haemagglutinin: adhesin: therapy: diagnosis: vaccine: antigen.

XX Porphyromonas gingivalis strain W50.

XX Key Location/Qualifiers

FT Peptide 1..227

FT /label= Pro-pro-peptide

FT Cleavage-site 227..228

FT Protein 228..719

FT /label= Prrr45

FT /note= "45 kDa Arg-specific thiol protease"

FT Cleavage-site 719..720

FT Protein 720..1138

FT /label= Prrr44

FT /note= "44 kDa adhesin"

FT Cleavage-site 1138..1139

FT Protein 1139..1273

FT /label= Prrr15

FT /note= "15 kDa adhesin"

FT Cleavage-site 1273..1274

FT Protein 1274..1431

FT /label= Prrr17

FT /note= "17 kDa adhesin"

FT Cleavage-site 1431..1432

FT Protein 1432..1706

FT /label= Prrr27

FT /note= "27 kDa adhesin"

XX W09716542-A1.

XX 09-MAY-1997.

XX 30-OCT-1996: 96MO-AU00673.

XX 30-OCT-1995: 95AU-0006275.

XX (UYME-) UNIV MELBOURNE.

XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.

XX Bhogal PS, Reynolds EC, Slakeski N:

XX WPI: 1997-272112/24.

XX N-PSDB: AAT78850.

XX New antigenic protein complex from Porphyromonas gingivalis -

XX comprising Arg- and Lys- specific thiol endo-peptidase(s), used in

XX the detection, prevention and treatment of periodontal disease

XX Example 1: Fig 8b: 68pp: English.

XX A Prrr-Prrk cell surface protein of Porphyromonas gingivalis (PG)

XX comprises a 300 kDa complex composed of a 45 kDa arginine-specific

XX thiol protease and 44, 15, 17 and 27 kDa adhesins encoded by the

XX prr gene (AAT78850), and a 14 kDa lysine-specific thiol protease

XX and 39, 15 and 44 kDa adhesins (see AAW24787) encoded by the prrk

XX gene (AAT78851). A claimed antigenic complex comprises at least one

XX multimeric protein complex of prr and prrk each containing at

XX least one adhesion domain, the complex having a mol. wt. of over 200

XX kDa, and preferably comprises all 9 proteins of the prr-prrk

XX complex (see also AAW24780-85). It can be used in a claimed

XX composition to elicit an immune response directed against PG, and

XX in a claimed method of reducing the prospect of PG infection and/or

XX severity of disease. Antibodies directed against the complex are

XX claimed for use in treating PG infection. Unlike whole PG cells or

XX other previously prepared antigens based on fimbriae or the

XX capsule, the prr-prrk complex or component parts are safe and

XX effective antigens.

XX Sequence 1706 AA:

Query Match 97.8%: Score 714: DB 18: Length 1706:

Best Local Similarity 97.8%: Pred No. 1.2e-64:

Matches 131: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

QY 1 ADETFESSSTHGEAPAEWTTIDADGEGWCLSSGQDMLTAHGTNVVSSFSWNGMA 60

DB 1139 ADETFESSSTHGEAPAEWTTIDADGEGWCLSSGQDMLTAHGTNVVSSFSWNGMA 1198

QY 61 LNPQNYLISKDVGATKRYYPVNDGFPDGHYAVMISKTGTNAGDPFTVFEETPGINK 120

DB 1199 LNPQNYLISKDVGATKRYYPVNDGFPDGHYAVMISKTGTNAGDPFTVFEETPGINK 1258

QY 121 GGARFGLSTEANGA 134

DB 1259 GGARFGLSTEANGA 1272

RESULT 12

AA96024

ID AA96024 standard: Protein: 439 AA.

XX AA96024:

AC AA96024:

DT 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin haga Harep4 product.

DE Haemagglutinin: haga: periodontal disease: vaccine: antibody:

KW Haemagglutinin: haga: periodontal disease: vaccine: antibody:

KW Harep4.

XX Porphyromonas gingivalis strain 381.

OS W09617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995: 95MO-US16108.

XX 09-DEC-1994: 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepine G, Patil JM, Progulskie-Fox A:

XX Tumwasorn S:

XX WPI: 1996-287181/29.

XX N-PSDB: AAT30648.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4: Page 114-115: 153pp: English.

XX Harep4 (AA96024) is the product of the Harep4 repeat unit (AAT30648)

XX of the haga gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin haga (see also AA96030). Harep4 and other

XX haga repeat unit products (see also AA96021-23) can be obtd. from

XX cells transformed with cells and used as vaccines to protect humans or

XX animals against periodontal disease. Expression in Salmonella

XX cells allows prodn. of live vaccine. Harep4-4 can also be used

XX to detect the presence of anti-P. gingivalis antibodies and to

XX raise monoclonal antibodies for diagnostic appn.

XX Sequence 439 AA:

QY Query Match 97.4%: Score 711: DB 17: Length 439:

Best Local Similarity 97.0%: Pred No. 3.7e-65:

Matches 130: Conservative 3: Mismatches 1: Indels 0: Gaps 0:

QY 1 ADETFESSSTHGEAPAEWTTIDADGEGWCLSSGQDMLTAHGTNVVSSFSWNGMA 60

```

Db      192  ADFTEFESSSTHGEAPAEWTTIDADGCGWCLSSGQLDMLTAHGCTNVVASFSGNMA 251
OY      61  LNPDNVYLISKDVTGATKVKYKYYVPVNDGFPGDHYAVMISKGTGTNAGDFTVFEETPNGINK 120
        |||||||
Db      252  LNPDNVYLISKDVTGATKVKYKYYAVNDGFPGDHYAVMISKGTGTNAGDFTVFEETPNGINK 311
        |||||||
OY      121  GGARFGLSTEANGA 134
        |||||||
Db      312  GGARFGLSTEADGA 325

RESULT 13
AAM69492 standard; Protein: 439 AA.
AAM69492:
AC      AAM69492:
XX
XX
XX      22-DEC-1998 (first entry)
DT
XX      Haemagglutinin protein ha9A, HAREP4.
DE
XX      Haemagglutinin protein; periodontal disease; vaccine; ha9A.
KM      Porphyromonas gingivalis.
OS      US5824791-A.
PN
XX      20-OCT-1998.
PD
XX      11-DEC-1995; 95US-0570311.
PE
XX      11-DEC-1995; 95US-0570311.
PR      11-DEC-1995; 95US-0570311.
PR      08-SEP-1988; 88US-0241640.
PR      25-JAN-1991; 91US-0647119.
PR      09-DEC-1994; 94US-0353485.
XX
XX      (UABR-) UAB RES FOUND.
PA      (UFL ) UNIV FLORIDA.
PI
XX      Han N, Lantz M, Lepine G, Patil JM, Proguliske-Fox A;
PI      Tummasorn S;
XX
XX      WPI: 1998-582627/49.
XX      N-PSDB; AAV58879.
XX
XX      Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX      and/or protease poly:peptide(s))
XX
XX      Claim 1; Column 139-144; 101pp; English.
XX
XX      This sequence is encoded by a Porphyromonas gingivalis gene of the
XX      invention. This sequence represents the ha9A haemagglutinin protein. The
XX      polypeptides are used to produce antibodies to organisms associated with
XX      periodontal disease. The antibodies are also used in purification and
XX      identification procedures. The genes and polypeptides are used as
XX      vaccines against periodontal disease.
SQ      Sequence 439 AA:

Query Match      97.4%; Score 711; DB 19; Length 439;
Best Local Similarity 97.0%; Pred. No. 3,7e-65;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Db      312  GGARFGLSTEADGA 325

RESULT 14
AAR96021 standard; Protein: 450 AA.
AAR96021:
AC      AAR96021:
XX
XX      04-SEP-1996 (first entry)
DT
XX      P. gingivalis haemagglutinin ha9A HAREP1 product.
DE
XX      Haemagglutinin; ha9A; periodontal disease; vaccine; antibody;
XX      HAREP1.
XX      Porphyromonas gingivalis strain 381.
XX      WO9617936-A2.
XX
XX      13-JUN-1996.
PD
XX      11-DEC-1995; 95WO-US16108.
PF
XX      09-DEC-1994; 94US-0353485.
PR
XX      (UABR-) UAB RES FOUND.
PA      (UFL ) UNIV FLORIDA.
PI
XX      Han N, Lantz M, Lepine G, Patil JM, Proguliske-Fox A;
PI      Tummasorn S;
XX
XX      WPI: 1996-287181/29.
XX      N-PSDB; AAT30645.
XX
XX      Porphyromonas gingivalis genes and proteins - used in the detection
XX      and vaccination against periodontal disease
XX
XX      Claim 4; Page 103-104; 153pp; English.
XX
XX      HAREP1 (AAR96021) is the product of the HAREP1 repeat unit (AAT30645)
XX      of the ha9A gene (AAT30654) of P. gingivalis 318. It forms part
XX      of haemagglutinin ha9A (see also AAR96030). HAREP1 and other
XX      ha9A repeat unit products (see also AAR96022-24) can be obtained from
XX      transformed host cells and used as vaccines to protect humans or
XX      animals against periodontal disease. Expression in Salmonella
XX      cells allows production of live vaccine. HAREP1-4 can also be used
XX      to detect the presence of anti-P. gingivalis antibodies and to
XX      raise monoclonal antibodies for diagnostic application.
SQ      Sequence 450 AA:

Query Match      97.4%; Score 711; DB 17; Length 450;
Best Local Similarity 97.0%; Pred. No. 3,9e-65;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

AC AAK69489;

XX 22-DEC-1998 (first entry)

XX DE Haemagglutinin protein hagg, HAREPI.

XX KM Haemagglutinin protein; periodontal disease; vaccine; hagg.

XX OS Porphyromonas gingivalis.

XX PN US5824791-A.

XX PD 20-OCT-1998.

XX PF 11-DEC-1995; 95US-0570311.

XX PR 11-DEC-1995; 95US-0570311.

XX PR 08-SEP-1988; 88US-0241640.

XX PR 25-JAN-1991; 91US-0647119.

XX PR 09-DEC-1994; 94US-0353485.

XX PA (UABR-) UAB RES FOUND.

XX PI (UYFL) UNIV FLORIDA.

XX PI Han N, Lantz M, Lepine G, Patil JM, Progujske-Fox A;

XX PI Tumworn S.

XX DR WPI; 1998-58267/49.

XX DR N-PSDB; AAV58876.

XX PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin

XX PT and/or protease poly(peptide(s))

XX PS Claim 1; Column 121-126; 101pp; English.

XX CC This sequence is encoded by a Porphyromonas gingivalis gene of the

XX CC invention. This sequence represents the hagg haemagglutinin protein. The

XX CC polypeptides are used to produce antibodies to organisms associated with

XX CC periodontal disease. The antibodies are also used in purification and

XX CC identification procedures. The genes and polypeptides are used as

XX CC vaccines against periodontal disease.

XX SQ Sequence 450 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Search completed: June 26, 2003, 00:51:42
 Job time: 47.2961 secs

Query Match 97.4%; Score 711; DB 19; Length 450;
 Best Local Similarity 97.0%; Pred. No. 3.9e-65;
 Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTTFESSTGGEAPAEWTTIDADGGEGWICLSSGQIDMTLAHGTNNVSSFSMNQMA 60
 |||||||
 Db 186 ADFTTFESSTGGEAPAEWTTIDADGGEGWICLSSGQIDMTLAHGTNNVSSFSMNQMA 245
 |||||||
 OY 61 LNPDMNLISKDVYTGATKVKYYPVNDGFGDHYAVMISKYGTNAGDFTVVEETPNCINK 120
 |||||||
 Db 246 LNPDMNLISKDVYTGATKVKYYPVNDGFGDHYAVMISKYGTNAGDFTVVEETPNCINK 305
 |||||||
 OY 121 GGARFGLSTEANGA 134
 |||||||
 Db 306 GGARFGLSTEADGA 319
 |||||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:09 ; Search time 19.7059 Seconds
(without alignments)
200.076 Million cell updates/sec

Title: US-09-980-370-6

Perfect score: 730

Sequence: 1 ADPFTFESSHGEAPAEWT.....PGINKGARFGLSTEANCA 134

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTOS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	716	98.1	1687	2 US-08-570-311-29	Sequence 29, App1
2	716	98.1	1704	3 US-08-336-308A-10	Sequence 10, App1
3	716	98.1	1704	4 US-08-822-324-6	Sequence 6, App1
4	716	98.1	1704	4 US-09-490-931-10	Sequence 10, App1
5	715	97.9	1732	2 US-08-570-311-10	Sequence 10, App1
6	715	97.9	1732	2 US-08-353-485-10	Sequence 10, App1
7	711	97.4	439	2 US-08-570-311-22	Sequence 22, App1
8	711	97.4	439	2 US-08-570-311-16	Sequence 16, App1
9	711	97.4	2628	2 US-08-570-311-14	Sequence 14, App1
10	709	97.1	1087	2 US-08-570-311-18	Sequence 18, App1
11	709	97.1	1087	2 US-08-353-485-8	Sequence 8, App1
12	709	97.1	1358	2 US-08-570-311-27	Sequence 27, App1
13	704	96.4	456	2 US-08-570-311-18	Sequence 18, App1
14	704	96.4	456	2 US-08-570-311-20	Sequence 20, App1
15	176.5	24.2	497	2 US-08-570-311-2	Sequence 2, App1
16	176.5	24.2	497	2 US-08-353-485-2	Sequence 2, App1
17	83.5	11.4	316	3 US-08-682-643-4	Sequence 4, App1
18	79.5	10.9	316	1 US-08-038-932B-1	Sequence 1, App1
19	79.5	10.9	316	1 US-08-656-349-1	Sequence 1, App1
20	79.5	10.9	316	4 US-09-104-623A-5	Sequence 5, App1
21	79.5	10.9	316	4 US-09-019-532-5	Sequence 5, App1
22	78	10.7	785	3 US-09-265-108-2	Sequence 2, App1
23	78	10.7	785	4 US-09-479-264-2	Sequence 2, App1
24	73.5	10.1	319	3 US-08-682-643-3	Sequence 3, App1
25	72.5	9.9	435	4 US-09-072-917A-9	Sequence 9, App1
26	72.5	9.9	265	3 US-08-483-857-8	Sequence 8, App1
27	72.5	9.9	617	4 US-09-188-930-303	Sequence 303, App1

28	71	9.7	236	3 US-08-886-269-2	Sequence 2, App1
29	71	9.7	236	3 US-09-167-647-2	Sequence 2, App1
30	71	9.7	756	2 US-08-560-398-2	Sequence 10, App1
31	69.5	9.5	2123	4 US-08-968-685A-10	Sequence 33, App1
32	69	9.5	271	3 US-08-968-563-33	Sequence 33, App1
33	69	9.5	271	3 US-08-969-683A-33	Sequence 14, App1
34	69	9.5	271	4 US-09-297-928-14	Sequence 17, App1
35	68	9.3	250	3 US-08-968-563-17	Sequence 13, App1
36	68	9.3	250	4 US-08-969-683A-17	Sequence 13, App1
37	68	9.3	250	4 US-09-297-928-13	Sequence 1, App1
38	68	9.3	288	2 US-08-875-062-1	Sequence 3, App1
39	68	9.3	338	1 US-08-606-888A-3	Sequence 7, App1
40	68	9.3	531	2 US-07-862-588B-7	Sequence 33, App1
41	67	9.2	2353	4 US-09-377-155-33	Sequence 4, App1
42	67	9.2	2353	4 US-08-913-942-4	Sequence 33, App1
43	67	9.2	2353	4 US-09-669-974-33	Sequence 33, App1
44	67	9.2	2354	4 US-09-268-347-47	Sequence 47, App1
45	67	9.2	2411	4 US-09-268-347-36	Sequence 36, App1

ALIGNMENTS

RESULT 1
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/570,311
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15_C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:

LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Query Match 98.1%; Score 716; DB 2; Length 1687;
Best Local Similarity 97.8%; Pred. No. 3.8e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTFESSSTHGEAPAEWTTTIDADDGEGWLCSSGQDMLTAHGTNNVSSFSNNGMA 60
DB 1120 ADFTFESSSTHGEAPAEWTTTIDADDGEGWLCSSGQDMLTAHGTNNVSSFSNNGMA 1179
QY 61 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNKINK 120
DB 1180 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNKINK 1239
QY 121 GGARGLSTEANGA 134
DB 1240 GGARGLSTEANGA 1253

RESULT 2

US-08-336-308A-10
Sequence 10, Application US/08336308A
Patent No. 6017532

GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 98.1%; Score 716; DB 3; Length 1704;

Best Local Similarity 97.8%; Pred. No. 3.8e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTFESSSTHGEAPAEWTTTIDADDGEGWLCSSGQDMLTAHGTNNVSSFSNNGMA 60
DB 1137 ADFTFESSSTHGEAPAEWTTTIDADDGEGWLCSSGQDMLTAHGTNNVSSFSNNGMA 1196
QY 61 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNKINK 120
DB 1197 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNKINK 1256
QY 121 GGARGLSTEANGA 134
DB 1257 GGARGLSTEANGA 1270

RESULT 3

US-08-822-324-6
Sequence 6, Application US/08822324
Patent No. 6129917

GENERAL INFORMATION:
APPLICANT: Potempa, Jan S.
APPLICANT: Travis, James
APPLICANT: Genco, Caroline A.
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,945
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 103-95 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 488-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-324-6

Query Match 98.1%; Score 716; DB 3; Length 1704;
Best Local Similarity 97.8%; Pred. No. 3.8e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTFESSSTHGEAPAEWTTTIDADDGEGWLCSSGQDMLTAHGTNNVSSFSNNGMA 60
DB 1137 ADFTFESSSTHGEAPAEWTTTIDADDGEGWLCSSGQDMLTAHGTNNVSSFSNNGMA 1196
QY 61 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNKINK 120
DB 1197 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNKINK 1256

OY 121 GGARFGLSTEANGA 134
DB 1257 GGARFGLSTEANGA 1270

RESULT 4

US-09-931-10
Sequence 10, Application US/09490931
Patent No. 62747718

GENERAL INFORMATION:

APPLICANT: Treavis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Phillip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,931
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/336,308
FILING DATE:

APPLICATION DATA:

APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 21-93C

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8080

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-931-10

Query Match 98.1%; Score 716; DB 4; Length 1704;
Best Local Similarity 97.8%; Pred. No. 3.8e-67;

Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTEFEESSTHGEPAEWTTTIDADGEGWLCSSGQDLMTAHGCTNVVSSFSNMGMA 60
DB 1137 ADFTEFEESSTHGEPAEWTTTIDADGEGWLCSSGQDLMTAHGCTNVVSSFSNMGMA 1196
OY 61 LNPBNYLSKDVATGATKVKYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFEEETPNGLINK 120
DB 1197 LNPBNYLSKDVATGATKVKYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFEEETPNGLINK 1256
OY 121 GGARFGLSTEANGA 134
DB 1257 GGARFGLSTEANGA 1270

RESULT 5

US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Progulsk-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patel, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. Whitlock
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 424
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: 0F15.C3

TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10

Query Match 97.9%; Score 715; DB 2; Length 1732;
Best Local Similarity 97.8%; Pred. No. 5e-67;

Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTEFEESSTHGEPAEWTTTIDADGEGWLCSSGQDLMTAHGCTNVVSSFSNMGMA 60
DB 1157 ADFTEFEESSTHGEPAEWTTTIDADGEGWLCSSGQDLMTAHGCTNVVSSFSNMGMA 1216
OY 61 LNPBNYLSKDVATGATKVKYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFEEETPNGLINK 120
DB 1217 LNPBNYLSKDVATGATKVKYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFEEETPNGLINK 1276
OY 121 GGARFGLSTEANGA 134
DB 1277 GGARFGLSTEANGA 1290

RESULT 6

US-08-353-485-10

Sequence 10, Application US/08353485

Patent No. 5830710

GENERAL INFORMATION:

APPLICANT: Progulsk-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guyalaine

APPLICANT: Han, Naiming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1732 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-353-485-10

Query Match 97.9%; Score 715; DB 2; Length 1732;

Best Local Similarity 97.8%; Pred. NO. 5e-67;

Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADEFTFESSYHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVSSFSWNGMA 60

DB 1157 ADEFTFESSYHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVSSFSWNGMA 1216

QY 61 LNPONTYLSKDYTGATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFETPNK 120

DB 1217 LNPONTYLSKDYTGATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFETPNK 1276

QY 121 GGARFGLSTENGGA 134

DB 1277 GGARFGLSTENGGA 1290

RESULT 7

US-08-570-311-22

Sequence 22, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Progulsk-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guyalaine

APPLICANT: Han, Naiming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 439 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-22

Query Match 97.4%; Score 711; DB 2; Length 439;

Best Local Similarity 97.0%; Pred. NO. 2e-67;

Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADEFTFESSYHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVSSFSWNGMA 60

DB 192 ADEFTFESSYHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVSSFSWNGMA 251

QY 61 LNPONTYLSKDYTGATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFETPNK 120

DB 252 LNPONTYLSKDYTGATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVFETPNK 311

QY 121 GGARFGLSTENGGA 134

DB 312 GGARFGLSTENGGA 325

RESULT 8

US-08-570-311-16

Sequence 16, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tummasorn, Somying
APPLICANT: Lepine, Guyelaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-16

Query Match 97.4%; Score 711; DB 2; Length 450;
Best Local Similarity 97.0%; Pred. No. 2e-67;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNNVVSFSWNGMA 60
DB 186 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNNVVSFSWNGMA 245
QY 61 LNPNDNLISKDYVGARKVKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 246 LNPNDNLISKDYVGARKVKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 305
QY 121 GGAFFGLSTEANGA 134
DB 306 GGAFFGLSTEADGA 319

RESULT 9
US-08-570-311-14
Sequence 14, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tummasorn, Somying
APPLICANT: Lepine, Guyelaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Query Match 97.4%; Score 711; DB 2; Length 2628;
Best Local Similarity 97.0%; Pred. No. 2.4e-66;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNNVVSFSWNGMA 60
DB 685 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNNVVSFSWNGMA 744
QY 61 LNPNDNLISKDYVGARKVKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 745 LNPNDNLISKDYVGARKVKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 804
QY 121 GGAFFGLSTEANGA 134
DB 805 GGAFFGLSTEADGA 818

RESULT 10
US-08-570-311-8
Sequence 8, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UPL5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-8

Query Match 97.1%; Score 709; DB 2; Length 1087;
Best Local Similarity 97.0%; Pred. No. 1.1e-66;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAGGTNVVSSFSWNGMA 60
DB 520 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAGGTNVVSSFSWNGMA 579
QY 61 LNPDNVLLSKDYTGATKTKKYYYPVNDGFGPDHYAVMISKTGTNAGDFTVVFETPNCINK 120
DB 580 LNPDNVLLSKDYTGATKTKKYYYPVNDGFGPDHYAVMISKTGTNAGDFTVVFETPNCINK 639
QY 121 GGARFGLSTEANGA 134
DB 640 GGARFGLSTEANGA 653

RESULT 11
US-08-353-485-8
Sequence 8, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UPL5.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Query Match 97.1%; Score 709; DB 2; Length 1087;
Best Local Similarity 97.0%; Pred. No. 1.1e-66;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAGGTNVVSSFSWNGMA 60
DB 520 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAGGTNVVSSFSWNGMA 579
QY 61 LNPDNVLLSKDYTGATKTKKYYYPVNDGFGPDHYAVMISKTGTNAGDFTVVFETPNCINK 120
DB 580 LNPDNVLLSKDYTGATKTKKYYYPVNDGFGPDHYAVMISKTGTNAGDFTVVFETPNCINK 639
QY 121 GGARFGLSTEANGA 134
DB 640 GGARFGLSTEANGA 653

RESULT 12
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming

```
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

Query Match 97.1% Score 709; DB 2; Length 1358;
Best Local Similarity 97.0%; Pred. No. 1.5e-66;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTTFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVVSSFSNMGA 60
DB 791 ADFTTFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVVSSFSNMGA 850
QY 61 LNPDPNYLSKDVYATATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNK 120
DB 851 LNPDPNYLSKDVYATATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNK 910
QY 121 GGARFGLSTEANGA 134
DB 911 GGARFGLSTEANGA 924
```

RESULT 13

```
US-08-570-311-18
Sequence 18, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naïming
```

```
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-18

Query Match 96.4% Score 704; DB 2; Length 456;
Best Local Similarity 96.3%; Pred. No. 1.1e-66;
Matches 129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTTFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVVSSFSNMGA 60
DB 192 ADFTTFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVVSSFSNMGA 251
QY 61 LNPDPNYLSKDVYATATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNK 120
DB 252 LNPDPNYLSKDVYATATYKYYYPVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNK 311
QY 121 GGARFGLSTEANGA 134
DB 312 GGARFGLSTEANGA 325
```

RESULT 14

```
US-08-570-311-20
Sequence 20, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naïming
```

```
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-20

Query Match          96.4%; Score 704; DB 2; Length 456;
Best Local Similarity 96.3%; Pred. No. 1.1e-66;
Matches 129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTETFEESTHGEAPAEWTTTIDADGSGWLCSSGQLDMLTAHGCTNVVSSFSNMGMA 60
DB 192 ADFTETFEESTHGEAPAEWTTTIDADGSGWLCSSGQLDMLTAHGCTNVVSSFSNMGMA 251
QY 61 LNPDMYLLSKDYTGATKRYKYYYPVNDGFPGDHVAVMISKTGTNAGDFTVVEETPNKINK 120
DB 252 LNPDMYLLSKDYTGATKRYKYYYPVNDGFPGDHVAVMISKTGTNAGDFTVVEETPNKINK 311
QY 121 GGARFGLSTEANGA 134
DB 312 GGARFGLSTEADGA 325

RESULT 15
US-08-570-311-2
: Sequence 2, Application US/08570311
: Patent No. 5824791
: GENERAL INFORMATION:
: APPLICANT: Proguliske-Fox, Ann
: APPLICANT: Tumwasorn, Somying
: APPLICANT: Lepline, Guyalaine
: APPLICANT: Han, Nalming
```

```
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-2

Query Match          24.2%; Score 176.5; DB 2; Length 497;
Best Local Similarity 38.1%; Pred. No. 1.2e-10;
Matches 45; Conservative 16; Mismatches 44; Indels 13; Gaps 5;
```

```
QY 5 ETEFESTHGEAPAEWTTTIDADGSGWLCSSGQLDMLTAHGCTNVVSSFSNMGMA 59
DB 235 ESFDYQT--LPNGWTMIDADGDGHNW--LSTINVTATHTGDDGAMFSKSWASGAKI 289
QY 60 ALNPDNYLLSKDYTGATKRYKYYYPVNDGFP--GDHVAVMISKTGTNAGDFTV-VEEET 114
DB 290 DLSPDNYLVTPKRYVPENGKLSYWSQYPTNEHGVLESTGNEANFTIKLEET 347
```

Search completed: June 26, 2003, 00:53:58
Job time : 19.7059 secs

us-09-980-370-6.rapb

Page 1

OM protein - protein search, using sw model

```
Run on:      June 26, 2003, 00:50:39 ; Search time 28.3765 seconds
              (without annotations)
```

510.977 million cell updates/sec

Title: US-09-980-370-6
Parfact score: 730

Sequence: 1 ADFTETFSSTHGAPAEWT.....PNGINKGARFGLSTEANGA 134

Scoring table: BLOSUM62

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

```
Database : Published_Applications_AA:*
```

```

1: /cgn2.6/ptodata1/pubppa/US08_NEW_PUB.pcp.*
2: /cgn2.6/ptodata1/pubppa/US08_NEW_PUB.pcp.*
3: /cgn2.6/ptodata1/pubppa/US06_NEW_PUB.pcp.*
4: /cgn2.6/ptodata1/pubppa/US06_NEW_PUB.pcp.*
5: /cgn2.6/ptodata1/pubppa/US06_PBOCBM.pcp.*
6: /cgn2.6/ptodata1/pubppa/US07_NEW_PUB.pcp.*
7: /cgn2.6/ptodata1/pubppa/US07_PBOCBM.pcp.*
8: /cgn2.6/ptodata1/pubppa/PT08_PBOCBM.pcp.*
9: /cgn2.6/ptodata1/pubppa/US09_PBOCBM.pcp.*
10: /cgn2.6/ptodata1/pubppa/US09_NEW_PUB.pcp.*
11: /cgn2.6/ptodata1/pubppa/US10_PBOCBM.pcp.*
12: /cgn2.6/ptodata1/pubppa/US10_NEW_PUB.pcp.*
13: /cgn2.6/ptodata1/pubppa/US06_PBOCBM.pcp.*
14: /cgn2.6/ptodata1/pubppa/US06_PBOCBM.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	72.5	9.9	617	9	US-09-866-050A-303	Sequence 303, App
2	70	9.6	377	10	US-09-784-877-2	Sequence 2, App1
3	70	9.6	377	10	US-09-811-913-15	Sequence 15, App
4	70	9.6	377	10	US-09-811-913-15	Sequence 15, App
5	70	9.6	377	10	US-09-811-913-15	Sequence 15, App
6	68	9.3	250	9	US-09-308-207-33	Sequence 33, App
7	68	9.3	250	9	US-09-308-207-17	Sequence 17, App
8	67.5	9.2	363	10	US-09-815-242-1323	Sequence 1323, App
9	67.5	9.2	363	10	US-09-815-242-1323	Sequence 1323, App
10	67	9.2	256	9	US-10-125-693-18	Sequence 18, App1
11	67	9.2	256	9	US-09-880-748-1745	Sequence 1745, App
12	66	9.2	235	10	US-09-197-862-33	Sequence 33, App1
13	66	9.0	231	9	US-10-112-527-3	Sequence 3, App1
14	66	9.0	231	9	US-10-112-527-2	Sequence 2, App1
15	66	9.0	251	9	US-10-112-527-1	Sequence 1, App1
16	66	9.0	283	9	US-10-111-531-26	Sequence 26, App1
17	66	9.0	334	9	US-10-111-531-68	Sequence 68, App
18	66	9.0	334	9	US-10-111-531-81	Sequence 81, App
19	66	9.0	334	9	US-10-111-531-90	Sequence 90, App
20	66	9.0	334	9	US-10-111-531-95	Sequence 95, App
21	66	9.0	334	9	US-10-111-531-108	Sequence 108, App

ALIGNMENTS

20	66	9	0	334	9	US-10-14-531-117	Sequence 116, App
21	66	9	0	334	9	US-10-14-531-126	Sequence 126, App
22	66	9	0	334	9	US-10-14-531-135	Sequence 135, App
23	66	9	0	334	9	US-10-14-531-144	Sequence 144, App
24	66	9	0	334	9	US-10-14-531-152	Sequence 152, App
25	66	9	0	334	9	US-10-14-531-163	Sequence 163, App
26	66	9	0	334	9	US-10-14-531-171	Sequence 171, App
27	66	9	0	334	9	US-10-14-531-180	Sequence 180, App
28	66	9	0	334	9	US-10-14-531-189	Sequence 189, App
29	66	9	0	334	9	US-10-14-531-198	Sequence 198, App
30	66	9	0	575	9	US-09-738-626-123	Sequence 426, App
31	66	9	0	631	9	US-10-12-880-16	Sequence 16, App
32	66	9	0	817	9	US-10-11-527-7	Sequence 4, App1
33	65.5	9	0	410	10	US-10-12-632-11	Sequence 31, App1
34	65.5	9	0	410	10	US-09-205-448-1	Sequence 10, App1
35	65.5	9	0	441	9	US-09-916-444-10	Sequence 10, App1
36	65	8	9	263	9	US-09-738-626-6513	Sequence 6513, App
37	65	8	9	304	9	US-09-738-626-6554	Sequence 6554, App
38	65	8	9	315	8	US-09-588-425-200	Sequence 200, App
39	65	8	9	315	8	US-09-588-425-200	Sequence 200, App
40	65	8	9	315	8	US-09-588-425-200	Sequence 200, App
41	65	8	9	315	8	US-10-141-090-200	Sequence 200, App
42	65	8	9	315	8	US-09-867-6094-3	Sequence 3, App1
43	65	8	9	327	9	US-09-738-626-3532	Sequence 3532, App
44	65	8	9	929	9	US-09-738-626-2144	Sequence 5144, App
45	65	8	9	946	9	US-09-953-260-4	Sequence 44, App1

```

: SOFTWARE:FastSeq for Windows Version 4.0
: SEQ ID NO: 303
: LENGTH: 617
: TYPE: prt
: ORGANISM: Mouse
US-09-866-050A-303

Query Match
Best Local Similarity 29.3% Pred No. 21,
Matches 22, Conservative 9, Mismatches 31, Indels 13, Gaps 3

QY 3 FTEFFSSHGANAETTTIDADGGDEWICLSSGDLDLTAHGCTWVSSS--WNGMA 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 FKFSFSSHWECCHQ-----SGSLPSWIVSMQTDADQLTLHFSSKTYATFSAAMSOPG 354
QY 61 LNPDN-----VLISK 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 LSPDTPMPVPTVSIQ 369

RESULT 2
US-09-784-877-2
: Sequence 2, Application US/09784877
: Patent No. US20010012836a1
: GENERAL INFORMATION:

```

Thu Jun 26 11:58:15 2003

us-09-980-370-6.rapb

Page 2

APPLICANT: Erding Hu
APPLICANT: Yuan Zhu
TITLE OF INVENTION: HUMAN HISTONE DEACETYLASE GENE HD4
FILE REFERENCE: GP-70516-C1
CURRENT APPLICATION NUMBER: US/09/784,877
CURRENT FILING DATE: 2001-02-16
PRIOR APPLICATION NUMBER: 09/138,151
PRIOR FILING DATE: 1998-08-21
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 377
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-784-877-2

Query Match
Best Local Similarity 23.5%; Score 70; DB 10; Length 377;
Matches 32; Conservative 13; Mismatches 45; Indels 46; Gaps 5;

DB 42 LTAHGCTNVVSSFSNMGAL-----NPDNYLISKDVT--GATKYKYY---YPVNDGFP 89
177 LDLHGGVDEDAFSTSKVMTVSLHKFSFGFPGTGDVSDVGLGKGRYTSVNVPIQDSIQ 236
QY 90 GDHY-----AVMISKTGNNAGDFTVVEETPNCINK----- 120
DB 237 DEKTYQICESVLKEYVOAFNPKAVYLDLADDTIADGPMCSFNMTPVGIKCLKTYLQWL 296
QY 121 -----GGARGLSTEA 131
DB 297 ATLILGGGYNLANTRA 312

RESULT 3

US-09-817-913-15
Sequence 15 Application US/09817913
Patent No. US20020061860A1
GENERAL INFORMATION:
APPLICANT: LI, Zumei
APPLICANT: Bonfils, Claire
TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
FILE REFERENCE: 106101,144
CURRENT APPLICATION NUMBER: US/09/817,913
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,157
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 15
LENGTH: 377
TYPE: PRT
ORGANISM: Human
US-09-817-913-15

Query Match
Best Local Similarity 23.5%; Score 70; DB 10; Length 377;
Matches 32; Conservative 13; Mismatches 45; Indels 46; Gaps 5;

DB 42 LTAHGCTNVVSSFSNMGAL-----NPDNYLISKDVT--GATKYKYY---YPVNDGFP 89
177 LDLHGGVDEDAFSTSKVMTVSLHKFSFGFPGTGDVSDVGLGKGRYTSVNVPIQDSIQ 236
QY 90 GDHY-----AVMISKTGNNAGDFTVVEETPNCINK----- 120
DB 237 DEKTYQICESVLKEYVOAFNPKAVYLDLADDTIADGPMCSFNMTPVGIKCLKTYLQWL 296
QY 121 -----GGARGLSTEA 131
DB 297 ATLILGGGYNLANTRA 312

RESULT 4
US-09-817-538-15
Sequence 15 Application US/09817538
Patent No. US20020137162A1
GENERAL INFORMATION:
APPLICANT: LI, Zumei
APPLICANT: Bonfils, Claire
TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
FILE REFERENCE: 106101,144
CURRENT APPLICATION NUMBER: US/09/817,538
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,157
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 15
LENGTH: 377
TYPE: PRT
ORGANISM: Human
US-09-817-538-15

Query Match
Best Local Similarity 23.5%; Score 70; DB 10; Length 377;
Matches 32; Conservative 13; Mismatches 45; Indels 46; Gaps 5;

DB 42 LTAHGCTNVVSSFSNMGAL-----NPDNYLISKDVT--GATKYKYY---YPVNDGFP 89
177 LDLHGGVDEDAFSTSKVMTVSLHKFSFGFPGTGDVSDVGLGKGRYTSVNVPIQDSIQ 236
QY 90 GDHY-----AVMISKTGNNAGDFTVVEETPNCINK----- 120
DB 237 DEKTYQICESVLKEYVOAFNPKAVYLDLADDTIADGPMCSFNMTPVGIKCLKTYLQWL 296
QY 121 -----GGARGLSTEA 131
DB 297 ATLILGGGYNLANTRA 312

RESULT 5

US-09-308-207-33
Sequence 33 Application US/09308207
Publication No. US2003002232A1
GENERAL INFORMATION:
APPLICANT: MARIA DIAZ-TORRES ET AL.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
PRODUCTION OF 1,3 PROPANEDIOL
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
1870 South Winton Road
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/308,207
FILING DATE: 13-May-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-Nov-1996
ATTORNEY/AGENT INFORMATION:
NAME: Glaislet, Debora
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2

[illegible]

```

: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: GPP2
: SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-308-207-17

Query Match          9.3%: Score 68; DB 9; Length 250;
Best Local Similarity 24.0%: Pred. NO. 21;
Matches 35; Conservative 21; Mismatches 58; Indels 32; Gaps 8;

OY 2 DF-TETFESSHGEAPAEWTTIDADGDEGMICLSSGQL---DMLTAHGCTNVVSSFSNM 57
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 68 DFANEYVNLKEALELPVYGEKSIIEPVAVKLCNALNLPKEKNAVATISGTRDMAOKPFE 127
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 58 GMAL-NPDNYLISKDV-TGATRKVY-----YPVNDGPPGDHYAVMISKTGTMACDF 107
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 128 HLGIRRPXYFTIANDVKGKDPHEPEYLTGRNGLGYPINEQDP-----SKSK----- 173
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 108 TVVEEFPNGINKGAR---FGLST 129
    ||||| | | | | | | | | | | | | | | | | | | | | | | |
DB 174 VVEFDAPAGIAAGKAAGCKITIGIAT 199
    ||||| | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-815-242-13233
: Sequence 13233, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes In
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011a
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13233
: LENGTH: 693
: TYPE: prt
: ORGANISM: Streptococcus pneumoniae
US-09-815-242-13233

Query Match          9.3%: Score 68; DB 10; Length 693;
Best Local Similarity 21.7%: Pred. NO. 74;
Matches 23; Conservative 14; Mismatches 29; Indels 40; Gaps 4;

OY 26 GDGEGWLCSSGOLDMLTAHGCTNVVSSFSNMGALNPDNYLISKDVTGATRKVYYPVN 85
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 616 GDVGWGHVYARGRGVDGMEAHGNSQI-----VRAVYVFLA 648
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```
OY      86 DCFPGDHVAMISKGTGTNAGDTVVF---EETPNGI-----NKG 121
Db      649 EMF---GATYVLRASASQGRGTGMVFDHYEDVPKSVQEEIIRKNG 691

RESULT 8
US-10-125-692-18
; Sequence 18, Application US/10125692
; Publication No. US2003004429A1
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan
; APPLICANT: Hayashi, Fumitaka
; APPLICANT: Smith, Kelly D.
; APPLICANT: Underhill, David M.
; APPLICANT: Ozlinsky, Adrian
; TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
; FILE REFERENCE: P-TS 5155
; CURRENT APPLICATION NUMBER: US/10/125,692
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/285,477
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 506
; TYPE: PRT
; ORGANISM: S. typhimurium2
US-10-125-692-18

Query Match          9.2%: Score 67.5; DB 9; Length 506;
Best Local Similarity 27.3%: Pred. No. 57;
Matches 36; Conservative 14; Mismatches 53; Indels 29; Gaps 6;

OY      20 TTIDADGDEGMLCLSSGQLDMLTAHGSTNVVSSFSMGMALNPN---YLISKDVTGAT 76
Db      196 TTIDVSGLDLDAI-----KATGCTNGTASVGTGAVKFDADNNKTFYIGFTGAD 246

OY      77 KVVY-YYPVNDGFGD-HYAMISKGTGNAGDT-----VFEETPNGINKG 122
Db      247 AANGDDEVAVVDGTVTLAAGATKTPMAGATTKEVQELKDPVAVVSADAKNALLGG 306

OY      123 ARFGLSTEANGA 134
Db      307 VD---ATDANGA 315

RESULT 9
US-09-880-748-1745
; Sequence 1745, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1745
; LENGTH: 256
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-880-748-1745

Query Match          9.2%: Score 67; DB 9; Length 256;
Best Local Similarity 19.7%: Pred. No. 28;
Matches 31; Conservative 21; Mismatches 49; Indels 56; Gaps 7;

OY      16 PAEWTTIDADGDEGMLCLSSGQL--DWLTAHGTNVVSSFSWNG----- 58
Db      14 PSETLSLCTCTSGS---ISSGNYMSVNRQHPG---KGLEWIGITIIDIGNTYNSLKS 66

OY      59 ---MALNDPNYLISKDVTGAT-----KVVYYPVNDGFPGDHY-----AV 95
Db      67 RVTMSVPTSKNQSFLELTSVTAADTAAYYCARVPYDTSGGLGEYYGDMVGQGTLY 126

OY      96 MISKGTNAGDFTVFEETPNGINKGARFGLSTEAN 132
Db      127 TVSSGGGSGG-----GSGGGGGSVLTQPAS 154

RESULT 10
US-09-797-862-33
; Sequence 33, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
; APPLICANT: PEAK, IAN RICHARD ANSELM
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: MOXON, E. RICHARD
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0134
; CURRENT APPLICATION NUMBER: US/09/797,862
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 2353
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-797-862-33

Query Match          9.2%: Score 67; DB 10; Length 2353;
Best Local Similarity 25.4%: Pred. No. 43e+02;
Matches 35; Conservative 12; Mismatches 55; Indels 36; Gaps 7;

OY      20 TTIDADGDE-----GWLCLSSG---OLDMLTAHGTNVVSSFSWNG 58
Db      1603 TVETDGDKEGNGLVAKAVIDAVNKAQMRVKTGGANGDDEFATVASGNTVFA-DGNG 1661

OY      59 MALNDPNYLISKDVTGATKVVYYPVNDGFPGCHYAMISKGTGNAGDFTVFEETPNGI 118
Db      1662 TTAE-----VTKANDSITVKKYVNVKVAADGLKDGKIDKADTTVTVADGKVT--ADN-- 1711

OY      119 NKGARF-----GLSTEAN 132
Db      1712 NGDGKKFVDASGLADALN 1729

RESULT 11
US-10-112-527-3
; Sequence 3, Application US/10112527
; Patent No. US20020168722A1
; GENERAL INFORMATION:
; APPLICANT: Grieninger, Gerd
; APPLICANT: Applegate, Dianne
; APPLICANT: Stoike-Steben, Iara
; TITLE OF INVENTION: No. US20020168722A1e1 Cleaved Fragments of Fibrinogen
; FILE REFERENCE: Docket 454-24 CON
; CURRENT APPLICATION NUMBER: US/10/112,527
; CURRENT FILING DATE: 2002-03-29
```

PRIOR APPLICATION NUMBER: US 09/373,157
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 221
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-527-3

Query Match
Best Local Similarity 22.9%; Score 66; DB 9; Length 221;
Matches 24; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

OY 2 DFTFFESTHGEAPAEWTTIDADGDEGWLCLSSGQDLMTAHGNTNVSSFSMNGMAL 61
DB 39 NFNRWODYKRG-----FGLNDEGEFEW--LGNDYLHLTGSGVLRLEEDMAG--- 88
OY 62 NPDNYLSKDVYTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 106
DB 89 -----NEAVAEYHFRV--GSEAEYALQVSSYEETAGD 119

RESULT 12
US-10-112-527-2
Sequence 2, Application US/10112527
Patent No. US20020168722A1
GENERAL INFORMATION:
APPLICANT: Griening, Gerd
APPLICANT: Applegate, Dianne
APPLICANT: Stolke-Steben, Lara
TITLE OF INVENTION: No. US20020168722A1 Cleaved Fragments of Fibrinogen
FILE REFERENCE: Docket 454-24 CON
CURRENT APPLICATION NUMBER: US/10/112,527
CURRENT FILING DATE: 2002-03-29
PRIOR APPLICATION NUMBER: US 09/373,157
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-527-2

Query Match
Best Local Similarity 9.0%; Score 66; DB 9; Length 236;
Matches 24; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

OY 2 DFTFFESTHGEAPAEWTTIDADGDEGWLCLSSGQDLMTAHGNTNVSSFSMNGMAL 61
DB 54 NFNRWODYKRG-----FGLNDEGEFEW--LGNDYLHLTGSGVLRLEEDMAG--- 103
OY 62 NPDNYLSKDVYTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 106
DB 104 -----NEAVAEYHFRV--GSEAEYALQVSSYEETAGD 134

RESULT 13
US-10-112-527-1
Sequence 1, Application US/10112527
Patent No. US20020168722A1
GENERAL INFORMATION:
APPLICANT: Griening, Gerd
APPLICANT: Applegate, Dianne
APPLICANT: Stolke-Steben, Lara
TITLE OF INVENTION: No. US20020168722A1 Cleaved Fragments of Fibrinogen
FILE REFERENCE: Docket 454-24 CON
CURRENT APPLICATION NUMBER: US/10/112,527
CURRENT FILING DATE: 2002-03-29
PRIOR APPLICATION NUMBER: US 09/373,157
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-527-1

Query Match
Best Local Similarity 9.0%; Score 66; DB 9; Length 251;
Matches 24; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

OY 2 DFTFFESTHGEAPAEWTTIDADGDEGWLCLSSGQDLMTAHGNTNVSSFSMNGMAL 61
DB 69 NFNRWODYKRG-----FGLNDEGEFEW--LGNDYLHLTGSGVLRLEEDMAG--- 118
OY 62 NPDNYLSKDVYTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 106
DB 119 -----NEAVAEYHFRV--GSEAEYALQVSSYEETAGD 149

RESULT 14
US-10-141-531-26
Sequence 26, Application US/10141531
Publication No. US20030100743A1
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helfetz, Peter
APPLICANT: Luginbuhl, Peter
APPLICANT: Muchhal, Umesh
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RET/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: Patentin version 3.1
SEQ ID NO 26
LENGTH: 283
TYPE: PRT
ORGANISM: Neurospora crassa
US-10-141-531-26

Query Match
Best Local Similarity 9.0%; Score 66; DB 9; Length 283;
Matches 27; Conservative 24; Mismatches 52; Indels 20; Gaps 5;

OY 12 HGEAPAEWTT--IDADGDEGWLCLSSGQDLMTAHGNTNVSSFSMNGMAL 58
DB 12 HEKVTVRFNVGVGVGDDGKLSHLYVKQVTTGKRETELEANLFYALGHDPATALYKQD 71
OY 59 MALNPDNYLSKDVYTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 112
DB 72 LETDADGCVVTKRGTTLTVEGVFAAGD-VQDKRYRQAITSAETGCMALDAEKFLSEHE 130

OY 113 ETP 115
DB 131 ETP 133

RESULT 15
US-10-141-531-68
Sequence 68, Application US/10141531
Publication No. US20030100743A1
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: del Val, Greg

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 26, 2003, 00:46:24 ; Search time 22.0706 Seconds
(without alignments) 583.673 Million cell updates/sec

Title: US-09-980-370-6

Perfect score: 730
Sequence: 1 ADFTETFEESTHGEAPAEWTTIDADGSGMCLSSQGLDMVLAHGTNNVSSFSNCGMA 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: PIR_73+
2: PIR_1+
3: PIR_2+
4: PIR_3+

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	716	98.1	1704	2 A55426	gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
2	715	97.9	1732	2 T30836	lysine-specific cysteine proteinase (EC 3.4.22.1) - Porphyromonas gingivalis
3	711	97.4	2628	2 T28651	hemagglutinin A - Porphyromonas gingivalis
4	706	96.7	1526	2 S48763	gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
5	90.5	12.4	218	2 B81094	isomerase, probable - Porphyromonas gingivalis
6	90.5	12.4	286	2 D75217	probable 2-acetylhydrolase - Porphyromonas gingivalis
7	89	12.2	418	2 T00154	hypothetical protein - Porphyromonas gingivalis
8	88.5	12.1	218	2 C81847	thermolysin (EC 3.4.22.1) - Porphyromonas gingivalis
9	83.5	11.4	552	2 A45564	thermolysin (EC 3.4.22.1) - Porphyromonas gingivalis
10	83	11.4	372	2 B70819	hypothetical protein - Porphyromonas gingivalis
11	82.5	11.3	326	2 A99283	hypothetical protein - Porphyromonas gingivalis
12	81.5	11.2	786	2 G69189	hypothetical protein - Porphyromonas gingivalis
13	80	11.0	428	2 H03761	hypothetical protein - Porphyromonas gingivalis
14	79.5	10.9	548	1 HYRSE	hypothetical protein - Porphyromonas gingivalis
15	78.5	10.8	531	2 B3706	hypothetical protein - Porphyromonas gingivalis
16	78.5	10.8	346	1 B4509	aldehyde lyase - Porphyromonas gingivalis
17	78.5	10.8	453	2 A45022	tryptophanase (EC 3.4.1.1) - Porphyromonas gingivalis
18	78.5	10.8	499	2 A27198	cellulase (EC 3.2.1.4) - Porphyromonas gingivalis
19	78	10.7	527	2 PNO114	microbial metalloprotease - Porphyromonas gingivalis
20	76	10.4	187	2 F64834	flammarin-like protein - Porphyromonas gingivalis
21	75.5	10.3	439	2 T02956	alpha-amylase (EC 3.2.1.3) - Porphyromonas gingivalis
22	75.5	10.3	607	2 S48528	endoxyanase - rum - Porphyromonas gingivalis
23	75.5	10.3	1034	2 A36108	integrin alpha-V - Porphyromonas gingivalis
24	75.5	10.3	759	2 A49398	cycloartenol synthase - Porphyromonas gingivalis
25	75	10.3	985	2 T28910	hypothetical protein - Porphyromonas gingivalis
26	75	10.2	479	2 A69832	probable acid-CoA lyase - Porphyromonas gingivalis
27	74.5	10.2	548	1 HYRSE	hypothetical protein - Porphyromonas gingivalis
28	74.5	10.2	550	1 T36746	probable serine/threonine protease - Porphyromonas gingivalis

30	74	10.1	453	2 B49022	tryptophanase (EC 3.4.1.1) - Porphyromonas gingivalis
31	74	10.1	504	2 S54744	cellulase (EC 3.2.1.4) - Porphyromonas gingivalis
32	74	10.1	354	2 G72858	hypothetical protein - Porphyromonas gingivalis
33	74	10.1	239	2 T28651	hemagglutinin A - Porphyromonas gingivalis
34	73.5	10.1	544	2 A42464	microbial metalloprotease - Porphyromonas gingivalis
35	73.5	10.1	546	2 J04113	neutral protease - Porphyromonas gingivalis
36	73.5	10.1	546	2 S72176	thermolysin (EC 3.4.22.1) - Porphyromonas gingivalis
37	73.5	10.1	546	2 S72176	thermolysin (EC 3.4.22.1) - Porphyromonas gingivalis
38	73.5	10.1	546	2 A81423	beta-glucosidase - Porphyromonas gingivalis
39	73.5	10.1	4936	2 AH2515	hypothetical protein - Porphyromonas gingivalis
40	73.5	10.1	435	2 S12625	alpha-amylase (EC 3.2.1.3) - Porphyromonas gingivalis
41	73	10.0	435	2 J07137	isochlorogenic acid synthase - Porphyromonas gingivalis
42	73	10.0	441	2 E84264	flagellar hook protein - Porphyromonas gingivalis
43	72.5	9.9	264	2 G69522	extracellular protease - Porphyromonas gingivalis
44	72.5	9.9	449	2 J07306	extracellular protease - Porphyromonas gingivalis
45	72.5	9.9	449	2 J07306	extracellular protease - Porphyromonas gingivalis

ALIGNMENTS

RESULT 1

A55426
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP
C:Species: Porphyromonas gingivalis
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
K:Accession: A55426; D53113
K:Protein: N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr, J.
J: Biol. Chem. 270, 1007-1010, 1995
A:Title: Molecular cloning and structural characterization of the Arg-gingipain protease from Porphyromonas gingivalis
A:Reference number: A55426; MIM:9518080; PMID:7536351
A:Accession: D53113
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-1704 <PAV>
A:Cross-references: GB:U15282; NID:9557067; PID:AAA69539.1; PID:9557068
R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
J: Biol. Chem. 269, 406-411, 1994
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso A:Reference number: A53113; MIM:94103245; PMID:8276827
A:Accession: D53113
A:Status: preliminary
A:Molecule type: protein
A:Residues: 228-249 <PIK>
A:Experimental source: H66
A:Note: sequence extracted from NCBI database (NCBI:141694)
C:Keywords: cysteine proteinase; hydrolase

Query Match

Best Local Similarity 98.1% Score 716; DB 2; Length 1704;
Best Local Similarity 97.8% Pred. No. 9.2e-58;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY	1	ADFTETFEESTHGEAPAEWTTIDADGSGMCLSSQGLDMVLAHGTNNVSSFSNCGMA	60
DB	1137	ADFTETFEESTHGEAPAEWTTIDADGSGMCLSSQGLDMVLAHGTNNVSSFSNCGMA	1196
QY	61	LNDPNTLISKVTATKRYKYPYPPNOCFGDHYRANISKTCGNMADCFVFEETPRGK	120
DB	1197	LNDPNTLISKVTATKRYKYPYPPNOCFGDHYRANISKTCGNMADCFVFEETPRGK	1256
QY	121	GSARFGLSTANCA	134
DB	1257	GSARFGLSTANCA	1270

RESULT 2

T30836
lysine-specific cysteine proteinase porphyrin (EC 3.4.22.-) - Porphyromonas gingivalis
N:Alternate names: lysine-specific cysteine proteinase 1, 60K
C:Species: Porphyromonas gingivalis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C:Accession: T30836; T30837; T30526; A53113

R:Barbey-Gallagher, C.A.; Han, N.; Patti, J.M.; Whitlock, J.; Progulake-Fox, A.; Lantz, J. Infect. 178: 2734-2741, 1996
 A:Title: Analysis of the prpC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
 A:Reference number: 220895; MUID:96213011; PMID:663659
 A:Accession: T30836
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1732 <BAR>
 A:Cross-references: EMBL:U42210; NID:91314325; PID:91314326; PIDN:AA06565.1
 R:Slakoski, N.; Cleal, S.M.; Reynolds, E.C.
 A:Submitted to the EMBL Data Library, October 1996
 A:Reference number: 220896
 A:Accession: T30837
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-795; 1,797-1289; 1,391-1478; 1,480-1732 <SLA>
 A:Cross-references: EMBL:073366; NID:92182811; PID:92182812; PIDN:AA060809.1
 R:Leal, J.F.; Leal, J.F.; Leal, J.F.; Leal, J.F.
 A:Submitted to the EMBL Data Library, October 1998
 A:Reference number: 220844; MUID:98298016; PMID:9625563
 A:Accession: T30526
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1350; 1,352-1363; 1,365-1447; 1,449-1732 <LEW>
 A:Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AA026523.1
 R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
 A:Biochem. J. 269: 406-411, 1994
 A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolation and characterization of the genes
 A:Reference number: A53113; MUID:9410345; PMID:8276827
 A:Accession: A53113
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 228-249 <PIK>
 A:Experimental source: H66
 A:Genetic source: H66
 A:Gene: prpC
 A:Keywords: cysteine proteinase; hydrolase
 Query Match 97.9%; Score 715; DB 2; Length 1732;
 Best Local Similarity 97.8%; Pred. No. 1.2e-57;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ADETFESSSTHGEAPAEWTTIDADGCGMCLSSGOLDMLTAHGTNNVVSFSWNGMA 60
 Db 1157 ADETFESSSTHGEAPAEWTTIDADGCGMCLSSGOLDMLTAHGTNNVVSFSWNGMA 1216
 Oy 61 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAGDFTVFEETPNCINK 120
 Db 1217 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAGDFTVFEETPNCINK 1276
 Oy 121 GGAARGLSTEAACA 134
 Db 1277 GGAARGLSTEAACA 1290

RESULT 3
 T28651
 hemagglutinin A - Porphyromonas gingivalis
 C:Species: Porphyromonas gingivalis
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Aug-2001
 C:Accession: T28651
 R:Han, N.; Whitlock, J.; Progulake-Fox, A.
 A:Infect. Immun. 64: 4000-4007, 1996
 A:Title: The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381 contains four
 A:Reference number: 220494; MUID:97047672; PMID:9826061
 A:Accession: T28651
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2628 <HAN>
 A:Cross-references: EMBL:U41807; NID:91552410; PID:91469916; PIDN:AA017128.1
 C:Genetics:

A:Gene: haga
 Query Match 97.4%; Score 711; DB 2; Length 2628;
 Best Local Similarity 97.0%; Pred. No. 4.4e-57;
 Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ADETFESSSTHGEAPAEWTTIDADGCGMCLSSGOLDMLTAHGTNNVVSFSWNGMA 60
 Db 685 ADETFESSSTHGEAPAEWTTIDADGCGMCLSSGOLDMLTAHGTNNVVSFSWNGMA 744
 Oy 61 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAGDFTVFEETPNCINK 120
 Db 745 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAGDFTVFEETPNCINK 804
 Oy 121 GGAARGLSTEAACA 134
 Db 805 GGAARGLSTEAACA 818

RESULT 4
 S49763
 gingipain B (PGC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)
 C:Species: Porphyromonas gingivalis
 C>Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997
 C:Accession: S49763
 R:Aduse-Opoku, J.; Muller, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.
 A:Submitted to the EMBL Data Library, November 1994
 A:Description: Cloning, sequence analysis and expression in Escherichia coli of prpR1
 A:Reference number: S49763
 A:Accession: S49763
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1526 <ADU>
 A:Cross-references: EMBL:X82680
 A:Genetic source: R1
 A:Gene: prpR1
 A:Keywords: cysteine proteinase; hydrolase
 Query Match 96.7%; Score 706; DB 2; Length 1526;
 Best Local Similarity 97.0%; Pred. No. 6.8e-57;
 Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 ADETFESSSTHGEAPAEWTTIDADGCGMCLSSGOLDMLTAHGTNNVVSFSWNGMA 60
 Db 1139 ADETFESSSTHGEAPAEWTTIDADGCGMCLSSGOLDMLTAHGTNNVVSFSWNGMA 1198
 Oy 61 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAGDFTVFEETPNCINK 120
 Db 1199 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAGDFTVFEETPNCINK 1258
 Oy 121 GGAARGLSTEAACA 134
 Db 1259 GGAARGLSTEAACA 1272

RESULT 5
 B81094
 isomerase, probable NMB138 [imported] - Neisseria meningitidis (strain MC58 serogroup
 C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: B81094
 R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.R.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; H. Qiu, H.; Yamashiro, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizzi, M.
 A:Science 287: 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappelli, R.
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: B81094
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <GT>
 A:Cross-references: GB:AE002482; GB:AE002098; NID:97226577; PIDN:AA01713.1; PID:9722

A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1338
C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 12.4%; Score 90.5; DB 2; Length 218;
Best Local Similarity 26.6%; Pred. No. 0.42;
Matches 38; Conservative 14; Mismatches 48; Indels 43; Gaps 6;

23 DADGDEG-----WLCSSGQDMLTAHG--GTNVVSFSNMG 59
Db 77 DSDGTGEGKDIICVAGYGVGLDITARDIQCRKEKGLPWLAKGFRHSACVSPFAAGR 136
Oy 60 ALNPDNYLISKDYTGATKVK-----YYPVNDGFPEDHYAVMISKT---GTNAGDFTVY 110
Db 137 IGNEPEKVLFLKONGVYLKQKRGDTGLMIYPIRE-----ILHKLADYGLGKGD--LV 185
Oy 111 FEETPNKINGKARFGLSTEANG 133
Db 186 FTGTSSGVGAIGAGDNLALDELGG 208

RESULT 6
D75217
probable 2-acetyl-1-alkylglycerophosph ocholone esterase PAB2176 - Pyrococcus abyssi (st
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: D75217
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: D75217
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <KAM>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49187.1; PID:el51508
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2176

Query Match 12.4%; Score 90.5; DB 2; Length 286;
Best Local Similarity 27.9%; Pred. No. 0.58;
Matches 29; Conservative 16; Mismatches 26; Indels 33; Gaps 5;

Oy 6 TFEESHTGAPAEATTIDADGDEGMLCSSLGQDMLTAHGNTNVVSFSMN----- 57
Db 102 TDFPRAHGESEGSKTIT--GDKE--ILDLSGAIDWLLSNTYTKITALIGFSMGAMVTR 156
Oy 58 -----GMALNPONVYLISKDYTGATKVKY-----YPV 84
Db 157 ALAEDERVCCGIADSPPIYI---DKTGAGLKTIFANLPFLYFI 197

RESULT 7
T00154
hypothetical protein 21 - Staphylococcus aureus phage phi PVL
C:Species: Staphylococcus aureus phage phi PVL
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T00154
R:Kaneko, Y.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.
Biosect. Biotechnol. Biochem. 61, 1960-1962, 1997
A:Title: Panton-Valentine leukocidin genes in a phage-like particle isolated from mitomy
A:Reference number: Z14119; MUID:98067870; PMID:9404084
A:Accession: T00154
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-418 <KAN>
A:Cross-references: EMBL:AB009666; NID:g3341907; PIDN:BA431894.1; PID:g3341928
C:Superfamily: Staphylococcus aureus phage phi PVL hypothetical protein 21

Query Match 12.2%; Score 89; DB 2; Length 418;
Best Local Similarity 27.3%; Pred. No. 1.2;

Matches 36; Conservative 19; Mismatches 55; Indels 22; Gaps 7;

Oy 20 TTIDADGDEGMLCSSLGQDMLTAHGNTNVVSFSNMG--MALNPDNYLISKDYTG 75
Db 94 TYIDGEGEDGG-----SSGITQWMDKTYSDSGKNGGTTISYGVVALVSDNNRVLLESTAS 149
Oy 76 TKVK-----YYYPVNDGFPG--DHYAVMISKT--GTNAGDFTVVEETPNKINGKARF-- 125
Db 150 SNIKSKQAPVYLYLPNDKVPGLNRFATLSNADNAYSSDGYIMFGSDENYDYGAGIRFSK 209
Oy 126 ----GLSTEANG 133
Db 210 ERNKGGLVQIYNG 221

RESULT 8
C81847
probable hydrolase NMA1552 [imported] - Neisseria meningitidis (strain Z2491 serogrou
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: C81847
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: C81847
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <PAR>
A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84779.1; PID:g738
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1552
C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 12.1%; Score 88.5; DB 2; Length 218;
Best Local Similarity 26.6%; Pred. No. 0.63;
Matches 38; Conservative 14; Mismatches 48; Indels 43; Gaps 6;

Oy 23 DADGDEG-----WLCSSGQDMLTAHG--GTNVVSFSNMG 59
Db 77 DSDGTGEGKDIICVAGYGVGLDITARDIQCRKEKGLPWLAKGFRHSACVSPFAAGR 136
Oy 60 ALNPDNYLISKDYTGATKVK-----YYPVNDGFPEDHYAVMISKT---GTNAGDFTVY 110
Db 137 IGNEPEKVLFLKONGVYLKQKRGDTGLMIYPIRE-----ILHKLADYGLGKGD--LV 185
Oy 111 FEETPNKINGKARFGLSTEANG 133
Db 186 FTGTSSGVGAIGAGDNLALDELGG 208

RESULT 9
A46564
thermolysin (EC 3.4.24.27) homolog precursor - Bacillus stearothermophilus
C:Species: Bacillus stearothermophilus
C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 22-Jun-1999
C:Accession: A46564
R:Kubo, M.; Inanaka, T.
J. Gen. Microbiol. 134, 1883-1892, 1988
A:Title: Cloning and nucleotide sequence of the highly thermostable neutral protease
A:Reference number: A46564; MUID:89235633; PMID:3149972
A:Accession: A46564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUB>
A:Cross-references: GB:M21663; NID:g143746; PIDN:AA802774.1; PID:g143747
A>Note: the authors translated the codon TCG for residue 32 as Thr, CTG for residue 5
C:Superfamily: thermolysin
C:Keywords: hydrolase; metalloprotease; zinc
F:378,382,402/Binding site: zinc (His, His, Glu) #status predicted
F:379,467/Active site: Glu, His #status predicted

Query Match 11.4% Score 83.5; DB 2; Length 552;
 Best Local Similarity 25.3%; Pred. No. 5.5;
 Matches 41; Conservative 28; Mismatches 52; Indels 41; Gaps 10;

QY 9 SSTH---GEAFAEWT---TIDADGGEGWCLSSGOLDMLTAHGTNVSSNNMGLN 62
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 338 SSVHISGQNNHFNNGSMVYGGDDGOTFIPLSGG-ID-VVAHELTHAVTDYR-AGLIYO 394
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 63 PDNYLSK---DYTGATKVKYKYPVNDG-----PG-----DHY 93
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 395 NESGAINEAISDIFG-FLVEFYANKNPDEIGEDVYTPGISDLSRMSDPKAYGDDPHY 453
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 94 AVNISKRTGNAGDF--YVFEETENGINKGAREGLSTEANG 133
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 454 SKRYGTQDMGVHNSIINKAKAYLSQGTHTYGVYSGIG 495
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 10

Probable carnitine operon oxidoreductase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: B70819
 R:Colt, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: B70819
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-373 <COL>
 A:Cross-references: GB:A022002; GB:AL123456; NID:93261544; PIDN:CA117610.1; PID:9326154
 C:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: fadE16
 C:Superfamily: acyl-CoA dehydrogenase

Query Match 11.4% Score 83; DB 2; Length 373;
 Best Local Similarity 28.3%; Pred. No. 3.9;
 Matches 39; Conservative 15; Mismatches 64; Indels 20; Gaps 6;

QY 9 SSTHGEAPAEWTTIDADGGEGWCLSSGOLDMLTAHGTN--VVSFSGMGMALNPVY 66
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 118 SSSHMAWY--STASADGG-----IAYRADKSWTTSAGFADYVVSAGSADGADVDLY 171
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 67 LSKDYTGATKVKYKYPVNDGFPDDHYAVMI-----SKGTNAGDTTVPFEETPNCI 118
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 172 AVPADTPGLRYAGTF--TGMLRLGNASAPMAVDIRIPDSYRLEAGGCGIMMOTVLPWF 229
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 119 NKGA--RFGSLTEANGA 134
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 230 NLGNNAVSLGLATATGA 247
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 11

Hypothetical protein dppC-2 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: A99283
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-aret, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.
 J. Mol. Biol. 261, 1-11, 1999
 A:Submitted to Genbank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: A99283
 A>Status: preliminary
 A:Molecule type: DNA
 A:Cross-references: NID:96598481; PIDN:AA69112.2; GSPDB:GN00139

A:Residues: 1-536 <RUB>
 A:Cross-references: GB:A006641; NID:913814480; PIDN:AA41520.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: dppC-2

Query Match 11.3% Score 82.5; DB 2; Length 536;
 Best Local Similarity 25.3%; Pred. No. 6.5;
 Matches 41; Conservative 22; Mismatches 34; Indels 65; Gaps 10;

QY 26 GGGEGWCLSSGOLDMLTAHGTNVSSFGMGMALNP-----NVLISKDYTA 75
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 164 GDEYFYIIQAS-----ISAY-ATN--AYTFNGYVINPSNATFWLFLAGNVLPTWLVLS 215
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 76 TVKKY-----YYPNDGFPDDHYAV-----MISKGTNAGD 106
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 216 TVFKYLGNGNNYILASSASAGETPMWTN--IPNESAVASYMLQSMFNSNTNYVE 273
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 107 FTVFEETPNCINKG-----GAREG-LSTEANGA 134
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 274 FTINY--IPNPNKLVYVYLSLYEFLGSRGVLTQDNNGA 313
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 12

Hypothetical protein MTH674 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: G69189
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadofora, R.; Vicalire, R.; Wang, Y.; Wlezbowski, J.; Gibson, R.; Jivan, K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: G69189
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-966 <MTH>
 A:Cross-references: GB:A0000847; GB:A000666; NID:92621756; PIDN:AA85179.1; PID:9262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH674

Query Match 11.2% Score 81.5; DB 2; Length 966;
 Best Local Similarity 36.1%; Pred. No. 16;
 Matches 26; Conservative 10; Mismatches 29; Indels 7; Gaps 3;

QY 68 ISKDYTGATKVKYKYPV-----NDGFPDDHYA-VMISKGTNAGDTTVPFEETPNCINKG 122
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 296 ISGYAMSLTLKLYTPVMDFTNPSPGELTISYTVGVGDDSDSLMLFLDTTPG--KGG 353
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 123 ARFGSLTEANGA 134
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 354 SYSGENTATGA 365
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 13

Cycloartenol synthase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H84481
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M. Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: H84481
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-759 <STO>
 A:Cross-references: GB:A002093; NID:96598481; PIDN:AA69112.2; GSPDB:GN00139

C:Genetics:
A:Gene: AT2907050

A:Map position: 2
C:Superfamily: yeast lanosterol synthase

Query Match 11.0%; Score 80; DB 2; Length 759;
Best Local Similarity 28.7%; Pred. No. 17;
Matches 25; Conservative 13; Mismatches 15; Indels 34; Gaps 7;

QY 13 GAPAEWTTIDADGEGWCLSSGQDMLTAHAGC-TNVVS-----SFSWNG-MA 60
DB 188 GGGP-----NDGGD-----MEGR-DWLNHGATNTSGMKWLSVLGAFEMSGNNP 235
QY 61 LNPBNYLISKDYTGATKKYKYYYPVNDG 87
DB 236 LPPEITWLT-----PYFLPIHPG 252

RESULT 14

503767

cellulase (EC 3.2.1.4) Z precursor - Erwinia chrysanthemi

N:Alternate names: endo-1,4-beta-glucanase; extracellular endoglucanase Z

C:Species: Erwinia chrysanthemi

A:Variety: strain 3937

C:Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 08-Oct-1999

C:Accession: S03767

R:Clustepol: A; Caml: B; Aymeric, J.L.; Ball, G.; Creuzet, N.

Mol: Microbiol. 2, 159-164, 1988

A:Title: Homology between endoglucanase Z of Erwinia chrysanthemi and endoglucanases of

A:Reference number: S03767; MUID:88216177; PMID:2835589

A:Accession: S03767

A:Molecule type: DNA

A:Residues: 1-428 <GUI>

A:Cross-references: EMBL:Y00540; NID:941091; PIND:CAA6604.1; PID:941092

C:Genetics:

A:Gene: celZ

C:Function: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce

A:Pathway: cellulose degradation

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-4/Domain: signal sequence #status predicted <SIG>

F:44-428/Product: cellulase Z #status predicted <MAT>

Query Match 10.9%; Score 79.5; DB 2; Length 428;
Best Local Similarity 28.7%; Pred. No. 9.5;
Matches 29; Conservative 15; Mismatches 26; Indels 31; Gaps 6;

QY 3 FTEFESTHGEA-----PAEWTTIDADGEGWCLSSGQD-WLTAH 45
DB 232 YLHFVAGHGESLNKKARQALNGLALFVTEMGTVNADGNGG---VNOTETDAWVTEM 287

QY 46 GGTNVSSFSNMGMALNDN-----YLISKDYTGATKKYK 79
DB 288 RDNNTIOLTONW---ALNDKNEGASTYTPDSKNLIESGKKVK 325

RESULT 15

HYBST

thermolysin (EC 3.4.24.27) precursor [validated] - Bacillus "thermoproteolyticus"

N:Alternate names: Bacillus thermoproteolyticus neutral proteinase; microbial metallopro

C:Species: Bacillus "thermoproteolyticus"

C:Date: 24-Apr-1984 #sequence_revision 07-Jun-1996 #text_change 15-Sep-2000

C:Accession: I40579; A00993; S41312

R:O'Donohue, M.J.; Roques, B.P.; Beaumont, A.

Biochem. J. 300, 599-603, 1994

A:Title: Cloning and expression in Bacillus subtilis of the npr gene from Bacillus therm

A:Reference number: I40579; MUID:94271180; PMID:8002967

A:Accession: I40579

A:Molecule type: DNA

A:Residues: 1-408, 'K', 410-548 <OAB>

A:Cross-references: EMBL:X76866; NID:9441266; PIND:CAA54291.1; PID:9441267

A:Note: submitted to the EMBL Data Library, December 1993

R:Itani, K.; Hermodson, M.A.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.

Nature New Biol. 238, 35-37, 1972

A:Title: Amino-acid sequence of thermolysin.

A:Reference number: A00993

A:Accession: A00993

A:Molecule type: Protein

A:Residues: 233-268, 'D', 270-350, 'E', 352-399, 'I', 401-548 <TIT>

R:Matthews, B.W.; Holmes, M.A.

submitted to the Brookhaven Protein Data Bank, February 1982

A:Reference number: A50631; PDB:1TLN

A:Contents: annotation: X-ray crystallography, 1.6 angstroms, residues 233-268, 'D', 27

J. Mol. Biol. 160, 623-639, 1982

A:Title: Structure of thermolysin redefined at 1.6 A resolution.

A:Reference number: A92888; MUID:83085812; PMID:7175940

A:Contents: annotation: X-ray crystallography, 1.6 angstroms

R:Matthews, B.W.; Weaver, L.H.; Kester, W.R.

J. Biol. Chem. 249, 8030-8044, 1974

A:Title: The conformation of thermolysin.

A:Reference number: A92165; MUID:75041142; PMID:4214815

A:Contents: annotation: X-ray crystallography, 2.3 angstroms

R:Burstein, Y.; Walsh, K.A.; Neurath, H.

Biochemistry 13, 205-210, 1974

A:Title: Evidence of an essential histidine residue in thermolysin.

A:Reference number: A90377; MUID:74052951; PMID:4808703

A:Contents: annotation: active site

C:Function:

A:Description: metalloendopeptidase preferentially hydrolyzes peptide bonds on the am

C:Superfamily: thermolysin

C:Keywords: calcium; extracellular protein; hydrolase; metalloproteinase; zinc

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-232/Domain: propeptide #status predicted <PRO>

F:233-548/Product: thermolysin #status experimental <MAT>

F:289,291,293/Binding site: calcium (Asp, Asp, Glu) #status experimental

F:370,409,417,419,422/Binding site: calcium (Asp, Asp, Glu, Asp, Glu, Glu) #status experim

F:374,378,398/Binding site: zinc (His, His, Glu) #status experimental

F:375,463/Active site: Glu, His #status predicted

F:409,413,417,422/Binding site: calcium (Glu, Asn, Asp, Glu) #status experimental

F:425,426,429,432/Binding site: calcium (Tyr, Thr, Ile, Asp) #status experimental

Query Match 10.9%; Score 79.5; DB 1; Length 548;
Best Local Similarity 24.7%; Pred. No. 13;
Matches 40; Conservative 28; Mismatches 53; Indels 41; Gaps 10;

QY 9 SSTRH---GEAPAWT---TIDADGEGWCLSSGQDMLTAHAGTNNVSSFSNMGMALN 62
DB 334 SSVHYSGVGNNAFNGSCQWVYGGDQGFPIPLSG-ID-VVAHELTHAVTDYT-AGLIYQ 390

QY 63 PDNVLISK---DVTGATKKYKYYYPVNDG-----PG-----DHY 93
DB 391 NESGALINEMSDIFG-TLVEFVANKRPWEIGEDVYTPDISDSLSMSDPAKYGDPDHY 449

QY 94 AVMISKGTNAGDF--TVVFETPNCINKGARFGLSTEANG 133
DB 450 SKRYTGTQDMGCVHINSGLINKAAVYLISOGTHYGVSVIG 491

Search completed: June 26, 2003, 00:52:22
Job time : 23.0706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:19 ; Search time 14.1882 Seconds

(without alignments)
391.721 Million cell updates/sec

Title: us-09-980-370-6

Perfect score: 730

Sequence: 1 ADLETFFESTHGEAPAEWT.....PENGKNGCARGLSTENGCA 134

Scoring table: BLAST62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SWISSPROT_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	714	97.8	989 1	PRTH_PORGI
2	711	97.4	2628 1	HAGA_PORGI
3	80	11.0	759 1	CASI_ARATH
4	79.5	10.9	316 1	THRE_BACTH
5	79.5	10.9	551 1	NPRE_BACST
6	78.5	10.8	348 1	GALM_STRTT
7	78.5	10.8	453 1	TNA1_SYMTT
8	78.5	10.7	426 1	GUN2_ERWCH
9	76	10.4	527 1	NPRE_BACBR
10	75.5	10.3	171 1	YCBV_ECOLI
11	75.5	10.3	1034 1	ITAV_CHICK
12	74.5	10.2	548 1	THRE_BACST
13	74	10.1	452 1	TNA2_SYMTT
14	74	10.1	504 1	GUNN_ERWCA
15	73.5	10.1	297 1	THW2_MYCTU
16	73.5	10.1	544 1	NPRE_BACCL
17	73	10.0	435 1	AM3D_ORISA
18	72.5	9.9	264 1	FLGG_BACSU
19	72	9.9	307 1	BLAC_MYCTU
20	72	9.9	519 1	VP6_MTV
21	72	9.9	1953 1	BIGA_SALTY
22	71.5	9.8	474 1	SHUJ_ECOLI
23	71	9.7	505 1	GUNN_ERWCA
24	71	9.7	519 1	VP6_MTV
25	70.5	9.7	607 1	XYNA_NDOPA
26	70	9.6	221 1	XYN2_EMENT
27	70	9.6	1045 1	GUNE_CELFT
28	70	9.6	1045 1	ITAV_HUMAN
29	69.5	9.5	312 1	PGTE_SALTY
30	69.5	9.5	547 1	FUMA_ECOLI
31	69.5	9.5	579 1	ITAV_SALTY
32	69.5	9.5	1044 1	ITAV_MOUSE
33	69.5	9.5	1176 1	CIAG_BACTU

34	69	9.5	249 1	GPPL_YEAST
35	69	9.5	463 1	STHA_PSEAE
36	69	9.5	705 1	CMBA_BACSU
37	69	9.5	958 1	TRH_DROME
38	68.5	9.4	321 1	CTSS_HOMAM
39	68.5	9.4	521 1	NPRE_BACSU
40	68.5	9.4	1048 1	ITAV_HUMAN
41	68.5	9.4	3672 1	LML2_CAEEL
42	68	9.3	250 1	GP2_YEAST
43	68	9.3	373 1	RF2_AQUAE
44	68	9.3	444 1	GUNN_ERWCA
45	68	9.3	523 1	TYD5_PAPSO

ALIGNMENTS

RESULT 1
ID PRTH_PORGI STANDARD; PRT; 989 AA.
AC P46071:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protease prth (EC 3.4.22.-).
GN PRTH.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M83;
RX MEDLINE=95012612; PubMed=7927685;
RA Fletcher H.M., Schenkein H.A., Macrina F.L.;
RT "Cloning and characterization of a new protease gene (prth) from
RT Porphyromonas gingivalis."
RL Infect. Immun. 62:4279-4286(1994).
RN [2]
RP ERRATUM.
RA Fletcher H.M., Schenkein H.A., Macrina F.L.;
RL Infect. Immun. 62:5707-5707(1994).
CC -1- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENABLE
CC P. GINGIVALIS TO EVADE COMPLEMENT-MEDIATED KILLING DURING THE
CC IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS
CC AND IS A VIRULENCE FACTOR.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch.)
CC
CC EMBL: U27483; AA051298.1; -
CC HSSP: P23882; IEMT.
CC DR MEMOPS; C25.001; -
CC DR InterPro: IPR001769; Peptidase_C25.
CC DR InterPro: IPR002376; formyl_transf.
CC DR Pfam: PF00551; formyl_transf. 1.
CC DR Pfam: PF01364; Peptidase_C25; 1.
CC KM HydroLase: Thiol protease; Repeat; Virulence.
CC FT REPEAT 528 323
CC FT REPEAT 581
CC SEQUENCE 989 AA; 110238 MW; F885FE8A3AC8944C CRC64;
Query Match 97.8%; Score 714; DB 1; Length 989;
Best Local Similarity 97.8%; Pred. No. 6.4e-60;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETPSSSTHGEAPAMETTTIDADGDEGMLSSGOLDMLTAHGTNNVSSFSMNMA 60
 DB 90 ADFTETPSSSTHGEAPAMETTTIDADGDEGMLSSGOLDMLTAHGTNNVSSFSMNMA 149
 OY 61 LNPDMVLSKDVGTGATKRYKYYYPVNDGPDGDAHVMISTGTMAGDFTVFEETPENGINK 120
 DB 150 LNPDMVLSKDVGTGATKRYKYYYPVNDGPDGDAHVMISTGTMAGDFTVFEETPENGINK 209
 OY 121 GGARFGLSTEANGA 134
 DB 210 GGARFGLSTEADGA 223

RESULT 2

HAGA_PORGI STANDARD: PRT: 2628 AA.
 AC 051845:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin A precursor.
 GN HAGA.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_Taxid=837;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RX MEDLINE=97047672; PubMed=89256061;
 RA Han N., Whitlock J., Prognaike-Fox A.;
 RT "The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381
 RL contains four large, contiguous, direct repeats."
 RL Infect. Immun. 64:4000-4007(1996).
 CC -1- FUNCTION: AGGLUTININATES ERYTHROCYTES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U01807; A017128.1;
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 6.
 KW Hemagglutinin; Virulence; Hydrolase; Thiol protease; Signal; Repeat.
 FT SIGNAL
 FT CHAIN 1
 FT DOMAIN 25 2628 HEMAGGLUTININ A.
 FT DOMAIN 25 539 PEPTIDASE C25-LIKE 1.
 FT DOMAIN 540 995 PEPTIDASE C25-LIKE 2.
 FT DOMAIN 996 1451 PEPTIDASE C25-LIKE 3.
 FT DOMAIN 1452 1907 PEPTIDASE C25-LIKE 4.
 FT DOMAIN 2074 2628 PEPTIDASE C25-LIKE 5.
 SO SEQUENCE 2628 AA; 283324 MW; 61C4DE32540C99DA CRC64;

Query Match 97.48; Score 711; DB 1; Length 2628;
 Best local Similarity 97.08; Pred. No. 3; e-59;
 Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETPSSSTHGEAPAMETTTIDADGDEGMLSSGOLDMLTAHGTNNVSSFSMNMA 60
 DB 685 ADFTETPSSSTHGEAPAMETTTIDADGDEGMLSSGOLDMLTAHGTNNVSSFSMNMA 744
 OY 61 LNPDMVLSKDVGTGATKRYKYYYPVNDGPDGDAHVMISTGTMAGDFTVFEETPENGINK 120
 DB 745 LNPDMVLSKDVGTGATKRYKYYYPVNDGPDGDAHVMISTGTMAGDFTVFEETPENGINK 804
 OY 121 GGARFGLSTEANGA 134
 DB 805 GGARFGLSTEADGA 818

RESULT 3

CAS1_ARATH STANDARD: PRT: 759 AA.

ID CAS1_ARATH
 AC P38605; P92967;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cycloartenol synthase (EC 5.4.99.8) (2,3-epoxysqualene-cycloartenol
 DE cyclase).
 GN CAS1 OR AT2G07050 OR T4E14.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX MEDLINE=94089686; PubMed=7505443;
 RA Corey E.J., Matsuda S.P.T., Bartel B.;
 RT "Isolation of an Arabidopsis thaliana gene encoding cycloartenol
 RT synthase by functional expression in a yeast mutant lacking
 RT lanosterol synthase by the use of a chromatographic screen."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11628-11632(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Bell C.R., Ketchum K.A., Lee J.J., Rongning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umeyam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cressy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768(1999).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shiozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Riken Arabidopsis full length cDNA clones (RAFTS) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC)."
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: (S)-2,3-epoxysqualene = cycloartenol.
 CC -1- SIMILARITY: BELONGS TO THE TERPENE CYCLASE/MUTASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 PFTB REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U02555; AAC04931.1;
 DR EMBL: AC005171; AA015015.1;
 DR EMBL: AY094394; AA019773.1;
 DR InterPro: IPR001330; Prenyltrans.
 DR InterPro: IPR002365; Terpene_synth.
 DR Pfam: PF00432; prenyltrans; 3.
 DR PROSITE: PS01074; TERPENE_SYNTHASES; 1.
 KW Isomerase; Repeat.
 FT REPEAT 147 188 PFTB 1.
 FT REPEAT 512 557 PFTB 2.
 FT REPEAT 589 629 PFTB 3.
 FT REPEAT 638 679 PFTB 4.
 FT REPEAT 700 741 PFTB 5.

FT CONFLICT 501 501 A -> E (IN REF. 1). C6C64;
 SQ SEQUENCE 759 AA: 8603 MW: DBA75CC37B1F74D C6C64;
 Query Match 11.0% Score 80; DB 1; Length 759;
 Best Local Similarity 28.7%; Pred. No. 4.8;
 Matches 25; Conservative 13; Mismatches 15; Indels 34; Gaps 7;
 OY 13 GEPAPMTTADGSGMGLSSQGLDWLTAAHG-TWVS-----SFSWNG-MA 60
 DB 188 GEP-----NDGCGD-----MEKGR-DWLHMGATNTSMGKMLSYLGAPEMGNP 235
 OY 61 LNPDWYLSKDVGTGATKYKYVPVNDG 87
 DB 236 LPPELWLT-----PYFLPIHNG 252

RESULT 4
 THER_BACTH STANDARD: PRT: 316 AA.
 AC P00800:
 DT 21-JUL-1986 (rel. 01, Created)
 DT 21-JUL-1986 (rel. 01, Last sequence update)
 DT 01-NOV-1997 (rel. 35, Last annotation update)
 DE Thermolysin (EC 3.4.24.27).
 OS Bacillus thermoproteolyticus.
 OC Bacillus; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1427;
 RN RP
 RA Tiliant K., Hermodson M.A., Ericsson L.H., Walsh K.A., Neurath H.;
 RT "Amino-acid sequence of thermolysin."
 RL Nature New Biol. 238:35-37(1972).
 RN RP
 RA Active Site:
 RX MEDLINE=74052951; PubMed=4808703;
 RT Bursstein Y., Walsh K.A., Neurath H.;
 RL "Evidence of an essential histidine residue in thermolysin."
 RL Biochemistry 13:205-210(1974).
 RN RP
 RA X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=83085812; PubMed=7175940;
 RT Holmes M.A., Matthews B.W.;
 RL "Structure of thermolysin refined at 1.6-A resolution."
 RL J. Mol. Biol. 160:623-639(1982).
 RN RP
 RA X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=75041142; PubMed=4214815;
 RT Matthews B.W., Weaver L.H., Kester W.R.;
 RL "The conformation of thermolysin."
 RL J. Biol. Chem. 249:8030-8044(1974).
 RN RP
 RA STRUCTURE BY NMR OF 255-316.
 RX MEDLINE=95086079; PubMed=7993910;
 RT Rico M., Jimenez M.A., Gonzalez C., de Filippis V., Fontana A.;
 RL "NMR solution structure of the C-terminal fragment 255-316 of
 thermolysin: a dimer formed by subunits having the native
 structure."
 RL Biochemistry 33:14834-14847(1994).
 RN RP
 RA STRUCTURE BY NMR OF 205-316.
 RX MEDLINE=97452621; PubMed=9305992;
 RT Conejero-Lara F., Gonzalez C., Jimenez M.A., Padmanabhan S.,
 RA Mateo P.L., Rico M.;
 RL "NMR solution structure of the 205-316 C-terminal fragment of
 thermolysin. An example of dimerization coupled to partial
 unfolding."
 RL Biochemistry 36:11975-11983(1997).
 CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Xaa-I-Lau > Xaa-I-Phe.
 CC -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS WHICH
 CC CONTRIBUTE TO THE UNUSUAL THERMAL STABILITY OF THIS ENDOPEPTIDASE.
 CC TWO ARE CLOSE TOGETHER AND PARTICIPATE IN A NETWORK OF SALT LINKS
 CC ALSO INVOLVING ASP-138, GLU-177, LYS-182, ASP-185, GLU-190, AND

CC ASP-191. A THIRD INTERACTS WITH ASP-57 AND A FOURTH WITH ASP-200.
 CC GUU AT POSITION 302 APPEARS TO FORM A SALT LINK WITH LYS-262.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 DR PIR: A00993; HYBST.
 DR PDB: 4TLN; 31-JAN-84.
 DR PDB: 5TLN; 31-JAN-84.
 DR PDB: 7TLN; 31-JAN-84.
 DR PDB: 8TLN; 30-APR-94.
 DR PDB: 17NN; 09-JAN-89.
 DR PDB: 27NN; 15-JAN-91.
 DR PDB: 3TNN; 09-JAN-89.
 DR PDB: 4TNN; 09-JAN-89.
 DR PDB: 5TNN; 09-JAN-89.
 DR PDB: 6TNN; 09-JAN-89.
 DR PDB: 7TNN; 15-OCT-94.
 DR PDB: 1THL; 31-JAN-84.
 DR PDB: 1TLF; 09-JAN-89.
 DR PDB: 1HVT; 31-JUL-94.
 DR PDB: 1TRL; 07-FEB-95.
 DR PDB: 1LNA; 08-MAY-95.
 DR PDB: 1LNB; 08-MAY-95.
 DR PDB: 1LNC; 08-MAY-95.
 DR PDB: 1LND; 08-MAY-95.
 DR PDB: 1LNE; 08-MAY-95.
 DR PDB: 1LNF; 08-MAY-95.
 DR MEROPS: M04.001;
 DR InterPro: IPR001570; Peptidase_M4.
 DR InterPro: IPR000130; Zn_MTPeptidase.
 DR Pfam: PF01447; Peptidase_M4_1.
 DR Pfam: PF02868; Peptidase_M4_C_1.
 DR PRINTS: PR00730; THERMOLYSIN.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR HydroLase: Metalloprotease; ZINC; 3D-structure.
 KM METAL 142 142
 FT ACT_SITE 143 143
 FT METAL 146 146
 FT METAL 166 166
 FT ACT_SITE 231 231
 FT STRAND 4 11
 FT TURN 13 14
 FT STRAND 17 24
 FT STRAND 28 29
 FT STRAND 31 32
 FT TURN 36 37
 FT STRAND 39 43
 FT TURN 45 46
 FT STRAND 53 54
 FT STRAND 56 57
 FT STRAND 61 62
 FT HELIX 65 88
 FT TURN 92 94
 FT TURN 100 106
 FT TURN 107 109
 FT STRAND 113 115
 FT STRAND 120 123
 FT STRAND 130 130
 FT HELIX 133 135
 FT HELIX 137 151
 FT TURN 152 153
 FT HELIX 159 179
 FT TURN 180 181
 FT STRAND 187 188
 FT TURN 192 192
 FT STRAND 193 193
 FT TURN 195 196
 FT TURN 198 199
 FT TURN 203 204
 FT TURN 208 211
 FT HELIX 212 212
 FT HELIX 217 219
 FT HELIX 225 229
 FT TURN 230 233

```

FT HELIX 234 246
FT STRAND 248 250
FT TURN 251 252
FT STRAND 253 255
FT HELIX 260 273
FT TURN 274 274
FT TURN 277 278
FT TURN 281 296
FT TURN 288 299
FT HELIX 301 312
FT TURN 313 314
SQ SEQUENCE 316 AA; 34333 MW; 004EDAF478744BB4 CRC64;

Query Match 10.98; Score 79.5; DB 1; Length 316;
Best Local Similarity 24.78; Pred. No. 2;
Matches 40; Conservative 28; Mismatches 53; Indels 41; Gaps 10;

OY 9 SSTR--GEAPAEWT--TIDADGDEGMCLSSGOLDMLTAGGTNNVSSFSNMGMLN 62
DB 102 SSVHYSQGYNNAFNGSEKYGDDGQDTFPLSG-ID-VVAHELTHAVDT-AGLIYQ 158
OY 63 PDNYLISK--DVTGATKVKYYPVNDGF-----PG-----DHY 93
DB 159 NESGAINEAISDFG-TLVEFYANKNPDMGEIDVYTPGISGDSLRSMSDPKYGDDHY 217
OY 94 AVMISKGTGNAGDF--TVVEETPNGINKGARGGLSTEANG 133
DB 218 SKRYTGQDNGGVHINSGIINKAAVLISGCGTHGVSVGIG 259

RESULT 5
NPRS_BACST
ID NPRS_BACST STANDARD: PRT: 551 AA.
AC P43133;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Bacillolysin precursor (EC 3.4.24.28) (Neutral protease).
GN NPRS OR NPRM.
OS Bacillus stearohermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
ON NCBI_TaxID=1422;
RX SEQUENCE FROM N.A., AND SEQUENCE OF 236-239.
RC STRAIN=TELNE;
RA MEDLINE=90368536; PubMed=2203733;
RA Mishiya Y., Imanaka T.;
RT "Cloning and nucleotide sequences of the bacillus stearohermophilus
RT neutral protease gene and its transcriptional activator gene.";
RL J. Bacteriol. 172:4861-4869(1990).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 236-240.
RC STRAIN=MK232;
RA MEDLINE=89235633; PubMed=3149972;
RA Kubo N., Imanaka T.;
RT "Cloning and nucleotide sequence of the highly thermostable neutral
RT protease gene from Bacillus stearohermophilus.";
RL J. Gen. Microbiol. 134:1883-1892(1988).
CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC -1- CATALYTIC ACTIVITY: Similar, but not identical, to that of
CC thermolysin.
CC -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in any way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: M34237; AAA2625.1; -
CC DR EMBL: M21663; AAB02774.1; -
CC DR HSSP: P00800; 1TRF.
CC DR MEROPS: M04.018; -
CC DR InterPro: IPR005075; pep_M4_propep.
CC DR InterPro: IPR001570; peptidase_M4.
CC DR InterPro: IPR000130; zn_mpeptide.
CC DR pfam: PF01447; Peptidase_M4.1.
CC DR pfam: PF02868; Peptidase_M4.C.1.
CC DR pfam: PF03813; pep_M4_propep.1.
CC DR PRINTS: PR00730; THERMOLYSIN.
CC DR PROSITE: PS00142; ZINC_PROTEASE.1.
CC DR HydroLase; Metalloprotease; zinc: Calcium; zymogen; signal.
CC FT SIGNAL 1 31
CC FT PROPEP 32 235
CC FT CHAIN 236 551
CC FT METAL 377 377
CC FT ACT_SITE 378 378
CC FT METAL 381 381
CC FT METAL 401 401
CC FT ACT_SITE 466 466
CC FT CONFICT 10 14
CC FT CONFICT 24 24
CC FT CONFICT 29 33
CC FT CONFICT 114 114
CC FT CONFICT 124 124
CC FT CONFICT 134 134
CC FT CONFICT 261 261
CC FT CONFICT 463 463
CC SQ SEQUENCE 551 AA; 60616 MW; FCF8B25A70129 CRC64;

Query Match 10.98; Score 79.5; DB 1; Length 551;
Best Local Similarity 24.78; Pred. No. 3.7;
Matches 40; Conservative 28; Mismatches 53; Indels 41; Gaps 10;

OY 9 SSTR--GEAPAEWT--TIDADGDEGMCLSSGOLDMLTAGGTNNVSSFSNMGMLN 62
DB 337 SSVHYSQGYNNAFNGSEKYGDDGQDTFPLSG-ID-VVAHELTHAVDT-AGLIYQ 393
OY 63 PDNYLISK--DVTGATKVKYYPVNDGF-----PG-----DHY 93
DB 394 NESGAINEAISDFG-TLVEFYANKNPDMGEIDVYTPGISGDSLRSMSDPKYGDDHY 452
OY 94 AVMISKGTGNAGDF--TVVEETPNGINKGARGGLSTEANG 133
DB 453 SKRYTGQDNGGVHINSGIINKAAVLISGCGTHGVSVGIG 494

RESULT 6
GALM_STRTR
ID GALM_STRTR STANDARD: PRT: 348 AA.
AC P21953;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aldose 1-epimerase (EC 5.1.3.3) (Mutarotase).
GN GALM.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CC -1- NCBI_TaxID=1308;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=A147;
CC RA MEDLINE=90298833; PubMed=1694527;
CC RA Poolman B., Royer T.J., Mairer S.E., Schmidt B.F.;
CC RT "Carbohydrate utilization in Streptococcus thermophilus:
CC characterization of the genes for aldose 1-epimerase (mutarotase) and
CC UDPglucose 4-epimerase.";
CC RL J. Bacteriol. 172:4037-4047(1990).
CC -1- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
CC is active on D-glucose, L-arabinose, D-xylose, D-galactose,

```

CC maltose and lactose (By similarity).
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose -> beta-D-glucose.
 CC -1- PATWAY: Hexose metabolism.
 CC -1- SIMILARITY: BELONGS TO THE ALDOSE EPIMERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC EMBL: M38175; AAA26945.1; -
 DR PIR: B44509; B44509.
 DR InterPro: IPR001823; Ald1-epimerase.
 DR Pfam: PF01263; Aldose-epim: 1.
 DR ProDom: PD006407; Ald1-epimerase: 1.
 DR PROSITE: PS00545; ALDOSE_1-EPIMERASE: 1.
 DR Isoemerase.
 DR ACT_SITE 180 180 POTENTIAL.
 SQ SEQUENCE 348 AA; 39012 MW; 17C3E027ABE1F92 CRC64;
 Query Match 10.8%; Score 78.5; DB 1; Length 348;
 Best Local Similarity 31.0%; Pred. No. 2.8; Mismatches 17; Gaps 3;
 Matches 26; Conservative 10; Indels 31; Indels 17; Gaps 3;
 OY 29 ERMGLSSGQDMTHAGCTNVSSFGNMGALNPNTLISKDVTGATKRYKYYVNDGFP 88
 DB 99 EGDNCL-----HGSPKMGQVQNNVVTMLNDYETK-----FIRLTSYDGF 142
 OY 89 PCDHAYMISYTGNGDFTVVE 112
 DB 143 PGD-VTVSISYRLNNRNLTLIFE 165
 RESULT 7
 TNA1_SYMTM STANDARD; PRT: 453 AA.
 AC P31013;
 DT 01-JUN-1993 (Rel. 26, Created)
 DT 01-JUN-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypophanase 1 (EC 4.1.99.1) (L-tryptophan indole-lyase 1) (Tnase 1).
 GN TNA1
 OS *Symbiobacterium thermophilum*.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Symbiobacterium.
 OX NCBI_TaxID:2734;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-36.
 RC STRAIN: IAM 13523; PubMed-1339259;
 RX MEDLINE-92384583; PubMed-1339259;
 RA Hirschner T., Swickl S., Morinouchi S., Beppu T.;
 RT "Cloning, nucleotide sequences, and overexpression in *Escherichia*
 RT coli of tandem copies of a tryptophanase gene in an obligately
 RT symbiotic thermophilic Symbiobacterium thermophilum.";
 RL Appl Environ Microbiol. 58:2633-2642(1992).
 CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> indole + pyruvate +
 CC NH(3)
 CC -1- COFACTOR: pyridoxal phosphate.
 CC -1- PATWAY: tryptophan catabolism.
 CC -1- SUBUNIT: homodimer (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC EMBL: AB010632; BAA24668.1; -

DR PIR: A49022; A49022.
 DR HSP: P31013; TRPL.
 DR InterPro: IPR001597; Beta-elim_lyase.
 DR Pfam: PF01212; Beta-elim_lyase: 1.
 DR ProDom: PD005927; Beta-elim_lyase: 1.
 DR PROSITE: PS00853; BETA-ELIM_LYASE: 1.
 DR Tryptophan catabolism: Lyase; PYRIDOXAL PHOSPHATE.
 DR BINDING 254 254 PYRIDOXAL PHOSPHATE (By similarity).
 SQ SEQUENCE 453 AA; 51012 MW; CAB332655C08812 CRC64;
 Query Match 10.8%; Score 78.5; DB 1; Length 453;
 Best Local Similarity 26.6%; Pred. No. 3.7; Mismatches 41; Indels 13; Gaps 4;
 Matches 25; Conservative 15; Indels 41; Indels 13; Gaps 4;
 OY 39 LDMTHAGCTNVSSFGNMGALNPNTLISKDVTGATKRYKYYVNDGFP 91
 DB 44 IDLLT-DSGTGMSGFQMSMMGDESYAGASVYRLKERTVDITGYEYVPHGRCAG 102
 OY 92 HYAMISYKGTAGDF--TVFEPTNGINKG 122
 DB 103 --KVAFSQILTRPGVYLSMFPDTRHGYQLG 134
 RESULT 8
 GUN2_ERWCH STANDARD; PRT: 426 AA.
 AC P07103;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Endoglucanase 2 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 2)
 DE (Cellulase 2) (EC2).
 GN CE1Z DP CE1S.
 OS *Erwinia chrysanthemi*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_TaxID:556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN: 3937;
 RX MEDLINE-86216177; PubMed-2835589;
 RA Gelsep A., Cami B., Aymeric J.-L., Ball G., Creuset N.;
 RT "Homology between endoglucanase 2 of *Erwinia chrysanthemi* and
 RT endoglucanase 2 of *Bacillus subtilis* and alkalophilic *Bacillus*.";
 RL Mol. Microbiol. 2:159-164(1988).
 RN [2]
 RP REVISIONS: AND DISULFIDE BOND.
 RC STRAIN: 3937;
 RX MEDLINE-94203057; PubMed-6152378;
 RA Bortoli-German I., Bryn E., Fry B., Chippaux M., Barras F.;
 RT Periplasmic disulphide bond formation is essential for cellulase
 RT secretion by the plant pathogen *Erwinia chrysanthemi*.";
 RL Mol. Microbiol. 11:545-553(1994).
 RN [3]
 RP MOTIVEMESIS AND DOMAINS.
 RX MEDLINE-91312860; PubMed-1677466;
 RA Fry B., Bortoli-German I., Halech J., Chippaux M., Barras F.;
 RT Cellulase Ec2 of *Erwinia chrysanthemi*: structural organization and
 RT importance of His98 and Glu133 residues for catalysis.";
 RL Protein Eng. 4:325-333(1991).
 RN [4]
 RP STEREOCHEMISTRY OF THE REACTION.
 RX MEDLINE-9222124; PubMed-156515;
 RA Barras F., Bortoli-German I., Bauman M., Kouvler J., Gey C.,
 RA Heyraud A., Hemmisse B.;
 RT Stereochemistry of the hydrolysis reaction catalyzed by
 RT endoglucanase 2 from *Erwinia chrysanthemi*.";
 RL FEBS Lett. 300:145-148(1992).
 RN [5]
 RP STRUCTURE BY NMR OF 365-426.
 RX MEDLINE-98070232; PubMed-9405041;
 RA Brun E., Mordard F., Gans P., Blackledge M.J., Barras F., Marion D.;
 RT "Solution structure of the cellulose-binding domain of the

Db 451 AAYLLAEGTHYGVVNCIG 470

RESULT 10

YCBV_ECOLI

ID YCBV_ECOLI STANDARD: PRT: 171 AA.

AC p75860:

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical fimbrial-like protein ycbv precursor.

GN YCBV OR B0943.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

ON NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655.

RX MEDLINE=97426617; PubMed=9278503;

RA Blatter F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.:

RT "The complete genome sequence of Escherichia coli K-12."

RL Science 277:1453-1474(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,

RA Kimura S., Kitegawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,

RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,

RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,

RA Yano M., Horikuchi T.:

RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 12.7-28.0 min region on the linkage map."

RL DNA Res. 3:137-155(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AEO00196; AAC74029.1; ALT_INIT.

DR EMBL: D90733; BAA35698.1; ALT_INIT.

DR Ecogene; EG13714; YCBV.

DR Interpro: IPR000259; Fimbrlal.

DR Pfam: PF00419; Fimbrlal; 1.

KW Hypothetical protein; Fimbrlal; Signal; Complete proteome.

KW SIGNAL.

FT CHAIN 1 17

FT CHAIN 18 171

FT CHAIN 18512 MW; 8E1CAF6A0AB5FA1 CRC64;

SO SEQUENCE

Query Match 10.3% Score 75.5; DB 1; Length 171;

Best Local Similarity 28.3% Pred. No. 2.4;

Matches 26; Conservative 15; Mismatches 44; Indels 7; Gaps 5;

Y 31 WCLSSGOLDW---LRAHGT-NVVSFSWNGMALNDNTLISKDVTGATKVKYYPVND 86

DB 7 WILFLG-LTWGELFLAHNDTNISGSFRNTCLADDSKOINVOLDVSLTRSHG-NY 64

OY 87 GFPGDHAAVMSKGTNAGDFTVFEETPGI 118

Db 65 G-PEKSEFIINLQDCGDIVSTVDYTFSCPTDGV 95

RESULT 11

ITAV_CHICK

ID ITAV_CHICK STANDARD: PRT: 1034 AA.

AC P26008:

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Integrin alpha-V precursor (Vltromectin receptor alpha subunit).

GN ITGAV.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

ON NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo.

RX MEDLINE=91104936; PubMed=1703004;

RA Bossy B., Reichardt L.F.:

RT "Chick integrin alpha V subunit molecular analysis reveals high

RT conservation of structural domains and association with multiple beta

RT subunits in embryo fibroblasts."

RL Biochemistry 29:10191-10198(1990)

CC -1- FUNCTION: THE ALPHA-V INTEGRINS ARE RECEPTORS FOR VITRONECTIN,

CC CYTOACTIN, FIBRONECTIN, FIBRINOGEN, LAMININ, MATRIX

CC METALLOPROTEINASE-2, OSTEOPOINTIN, THROMBOSPONDIN AND

CC VON WILLEBRAND FACTOR. THEY RECOGNIZE THE SEQUENCE R-G-D IN A WIDE

CC ARRAY OF LIGANDS.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA

CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A

CC DISULFIDE BOND. ALPHA-V ASSOCIATES WITH EITHER BETA-1, BETA-3,

CC BETA-5, BETA-6 OR BETA-8.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M60517; AAA49138.1; -.

DR PIR: A36108; A36108.

DR HSSP: P06756; IJY2.

DR Interpro: IPR000413; Integrin_alpha.

DR Pfam: PF00357; Integrin_A; 1.

DR Pfam: PF01839; FG-GAP; 5.

DR PRINTS: P01185; INTEGRINA.

DR SMART: SM00191; Int_alpha; 5.

DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW SIGNAL.

FT CHAIN 1 19

FT CHAIN 20 1034

FT CHAIN 20 875

FT CHAIN 877 1034

FT CHAIN 20 978

FT DOMAIN 979 1002

FT TRANSMEM 1003 1034

FT DOMAIN 1003 1034

FT REPEAT 35 96

FT REPEAT 101 166

FT REPEAT 167 224

FT REPEAT 237 290

FT REPEAT 291 350

FT REPEAT 355 415

FT REPEAT 420 473

FT CA_BIND 248 256

FT CA_BIND 302 310

FT CA_BIND 367 375

FT CA_BIND 431 439

FT SITE 1005 1009

FT DISULFID 77 85

POTENTIAL.

INTEGRIN ALPHA-V.

INTEGRIN ALPHA-V HEAVY CHAIN (POTENTIAL).

INTEGRIN ALPHA-V LIGHT CHAIN (POTENTIAL).

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

FG-GAP 1.

FG-GAP 2.

FG-GAP 3.

FG-GAP 4.

FG-GAP 5.

FG-GAP 6.

FG-GAP 7.

POTENTIAL.

POTENTIAL.

POTENTIAL.

POTENTIAL.

GFGR MOTIF.

BY SIMILARITY.

```
FT DISULFID 126 146 BY SIMILARITY.
FT DISULFID 160 173 BY SIMILARITY.
FT DISULFID 479 488 BY SIMILARITY.
FT DISULFID 494 551 BY SIMILARITY.
FT DISULFID 612 618 BY SIMILARITY.
FT DISULFID 684 697 BY SIMILARITY.
FT DISULFID 838 890 INTERCHAIN (BY SIMILARITY).
FT DISULFID 895 900 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 690 690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 860 860 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 931 931 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 966 966 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1034 AA; 114388 MW; D76B0B8A692DC684 CRC64;

Query Match 10.3%; Score 75.5; DB 1; Length 1034;
Best Local Similarity 23.8%; Pred. No. 18;
Matches 36; Conservative 21; Mismatches 47; Indels 47; Gaps 9;

OY 7 FESSTHGE-ANAEMTTTADGEGWLCSSQGLDLTAHGTNNV-----SFSNMGMAL 61
DB 149 FDGSKSVYAPACRSYTTIDADGCG---FCGGFSIDFTK---GDRVLLGGPSFYMOGOLI 202
OY 62 NP-----DNVLISKDYTGATKYKYYPVNDGFPDHYAVAMISKTGTN 103
DB 203 SDRAVELAKKDSKYSTRKYDQLATPRASAA-----FDDSYLG--YSVAV----- 246
OY 104 AGDFT-VVEETPNKNGKARFGLSTENG 133
DB 247 -GDFSGDIEDFVSGVPRARFLGMVSIYNG 276

RESULT 12
THER_BACST
ID THER_BACST STANDARD: PRT: 548 AA.
AC P06874;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Thermolysin precursor (EC 3.4.24.27) (Thermolysin neutral
  protease).
GN NPRT.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 230-243.
RX MEDLINE=85289021; PubMed=2933245;
RA Takagi M., Imanaka T., Alba S.;
RT "Nucleotide sequence and promoter region for the neutral protease
  gene from Bacillus stearothermophilus.";
RL J. Bacteriol. 163:824-831(1985).
CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Xaa-I-Leu > Xaa-I-Phe.
CC -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL: M11446; AAA22621.1;
DR PIR: A24824; HYBS.
DR HSSP: P00800; IHVT.
DR MEROPS: M04.001; -.
DR InterPro: IPR005075; Pep_M4_propep.
DR InterPro: IPR001570; Peptidase_M4.
DR InterPro: IPR000130; Zn_mtppepde.
DR Pfam: PF01447; Peptidase_M4_1.
DR Pfam: PF02866; Peptidase_M4_C1.
DR Pfam: PF03413; Pep_M4_propep; 1.
DR PRINTS: PR00730; THERMOLYSIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal.
FT SIGNAL 1 25
FT PROPEP 26 229 ACTIVATION PEPTIDE.
FT CHAIN 230 548 THERMOLYSIN.
FT METAL 374 374 ZINC (CATALYTIC).
FT ACT_SITE 375 375
FT METAL 378 378 ZINC (CATALYTIC).
FT METAL 398 398 ZINC (CATALYTIC).
FT ACT_SITE 463 463 PROTON DONOR.
SQ SEQUENCE 548 AA; 59580 MW; 5B4D2C1D7910539D CRC64;

Query Match 10.2%; Score 74.5; DB 1; Length 548;
Best Local Similarity 23.5%; Pred. No. 11;
Matches 38; Conservative 28; Mismatches 55; Indels 41; Gaps 10;

OY 9 GSTH---GEAPAEWT---TIDADGEGWLCSSQGLDLTAHGTNNVSSFSNMGMALN 62
DB 334 STVHYGRGNNAFNGSQMVYGDGQTFLEPFGS-G-ID-VVGHDLTHAVTYT-AGLVYQ 390
OY 63 PDNLISK---DVGATKVKYKYYPVNDG-----G-----DHY 93
DB 391 NESGAINEAMSDIFG-TLVEFYANRPWEIGEDITYTGAVAGDALRMSDPAKYGDPRHY 449
OY 94 AVMISKGTGTNAGDFT--VVEETPNKNGKARFGLSTENG 133
DB 450 SKRYTGTQDNGCVHTNSGIINKAAVLLSQGVHYGVSVNGIG 491

RESULT 13
TNA2_SYMTH
ID TNA2_SYMTH STANDARD: PRT: 452 AA.
AC P31015; G9R5M8;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanase 2 (EC 4.1.99.1) (L-tryptophan indole-lyase 2) (Tnase 2).
GN TNA2.
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Actinobacteriales (class); Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-34.
RX STRAIN=IAM 13621;
RX MEDLINE=92384583; PubMed=1339259;
RA Hirahara T., Suzuki S., Horinouchi S., Beppu T.;
RT "Cloning, nucleotide sequences, and overexpression in Escherichia
  coli of tandem copies of a tryptophanase gene in an obligately
  symbiotic thermophile, Symbiobacterium thermophilum.";
RL Appl. Environ. Microbiol. 58:2633-2642(1992).
RN [2]
RP SEQUENCE OF 1-35.
RX STRAIN=T;
RX MEDLINE=92144113; PubMed=1368766;
RX Suzuki S., Hirahara T., Horinouchi S., Beppu T.;
RT "Purification and properties of thermostable tryptophanase from an
  obligately symbiotic thermophile, Symbiobacterium thermophilum.";
RL Agric. Biol. Chem. 55:3059-3066(1991).
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate.
```


Thu Jun 26 11:58:17 2003

us-09-980-370-6.rspt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:47:39 ; Search time 40.9882 Seconds
(without alignments)
673.616 Million cell updates/sec

Title: US-09-980-370-6

Perfect score: 730
Sequence: 1 ADPTEFESSHGEAPAEWT.....PNCINKGCAFGCLSTANCA 134

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: SP-ATCHEP:21.*
2: SP-BACTERIA:.*
3: SP-FUNG1:.*
4: SP-HUMAN:.*
5: SP-INVERTEBRATE:.*
6: SP-MAMMAL:.*
7: SP-PLANT:.*
8: SP-VERTEBRATE:.*
9: SP-VERTEBRATE:.*
10: SP-VERTEBRATE:.*
11: SP-VERTEBRATE:.*
12: SP-VERTEBRATE:.*
13: SP-VERTEBRATE:.*
14: SP-VERTEBRATE:.*
15: SP-VERTEBRATE:.*
16: SP-VERTEBRATE:.*
17: SP-VERTEBRATE:.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	715	98.1	1223	092NB5	092NB5 porphyromon
2	715	98.1	1687	092NB5	092NB5 porphyromon
3	715	98.1	1704	051816	051816 porphyromon
4	715	97.9	1097	P72196	P72196 porphyromon
5	715	97.9	1732	052050	052050 porphyromon
6	715	97.9	1732	051817	051817 porphyromon
7	715	97.9	1732	007442	007442 porphyromon
8	714	97.8	1706	051839	051839 porphyromon
9	714	97.4	1723	P72197	P72197 porphyromon
10	709	97.1	1358	P66967	P66967 porphyromon
11	709	97.1	1723	P72194	P72194 porphyromon
12	706	96.7	1706	051838	051838 porphyromon
13	205	28.1	925	09FA40	09FA40 porphyromon
14	180	24.7	312	09K1B3	09K1B3 porphyromon
15	90.5	12.4	218	09J215	09J215 neisseria m
16	90.5	12.4	286	09V211	09V211 pyrococcus

17	89	12.2	418	9	080060	080060 staphylococ
18	89	12.2	1235	9	085DK2	085DK2 staphylococ
19	89	12.2	1261	9	09MBN9	09MBN9 staphylococ
20	88.5	12.1	9210	16	09J010	09J010 neisseria m
21	88.5	12.1	1217	16	08Y378	08Y378 neisseria s
22	87	11.9	910	3	P87211	P87211 orf10myces
23	85.5	11.7	2275	17	08T8E2	08T8E2 methanosarc
24	83.5	11.4	256	17	08TRD2	08TRD2 methanosarc
25	83.5	11.4	548	2	045779	045779 bacillus th
26	83.5	11.4	503	10	09M653	09M653 polyonatum
27	83	11.4	373	16	053926	053926 mycobacteri
28	83	11.4	806	3	09P478	09P478 agaricus bi
29	82.5	11.3	536	17	097YF2	097YF2 sulfolobus
30	81.5	11.2	454	2	086039	086039 alteromonas
31	81.5	11.2	966	17	026770	026770 methanobact
32	81	11.1	1372	2	054151	054151 shigella fl
33	81	11.1	534	2	09AL58	09AL58 shigella fl
34	79.5	10.9	735	16	08XNF9	08XNF9 lymnaea sta
35	79.5	10.8	348	2	08VS30	08VS30 streptococc
36	78.5	10.8	348	2	08VS84	08VS84 streptococc
37	78.5	10.7	499	2	045532	045532 bacillus su
38	78	10.7	785	3	09PBH5	09PBH5 humicola in
39	78	10.7	803	10	08W075	08W075 sorghum bic
40	78	10.7	1606	12	091LD6	091LD6 white spot
41	77.5	10.6	1039	5	09Y0B1	09Y0B1 dirosophila
42	77.5	10.5	698	2	09KH03	09KH03 alcaligenes
43	76	10.4	344	16	09BNM4	09BNM4 rhizobium j
44	76	10.4	680	2	052644	052644 rumioccocu
45	76	10.4	680	2	052644	052644 rumioccocu

ALIGNMENTS

RESULT 1

ID 092NB5 PRELIMINARY: PRT: 1223 AA.

AC 092NB5: 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 130K-HMGD (Fragment).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria: CPB group: Bacteroidetes; Bacteroidales: Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
FN [1]
PP SEQUENCE FROM N.A.

RC MEDJIN=9914316; Pubmed=998746;
RX Shibata Y., Hayekawa M., Takiguchi H., Shiroza T., Abiko Y.:
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products".
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL: AB019363; BAA34341.1; .
DR InterPro: IPR000977; DNA_119ase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 1223 AA: 131542 MW: 00225CD2BA9P91B3 CRC64;

Query Match 98.1%; Score 716; DB 2; Length 1223;
Best local similarity 97.8%; Pred. No. 1.2e-54;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADPTEFESSHGEAPAEWTIDAGDGCWGLSSGQIDWLTAGCTNVVSPFSMNCA 60
DB 648 ADPTEFESSHGEAPAEWTIDAGDGCWGLSSGQIDWLTAGCTNVVSPFSMNCA 707
QY 61 LNPENYLSKDVGTATGATKVKYYPVNDGFPGDHYAVMISGTGNGDPTVVEETPNCINK 120
|||||

```

Db 708 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 767
Qy 121 GGARFGLSTENGA 134
    |||||||
Db 768 GGARFGLSTENGA 781

RESULT 2
ID Q9R9B7 PRELIMINARY: PRT: 1687 AA.
AC Q9R9B7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hemagglutinin/protease.
GN HAGE.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriis; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA Han N., Dong H., Progulskie-Fox A.;
RT "Cloning and characterization of hage from P. gingivalis 381."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF028948; AAD01810.1;
DR HSSP: P95493; ICVR.
DR MEROPS: C25.001;
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1687 AA: 183702 MW: D085851A399FE70 CRC64:

Query Match
Best Local Similarity 98.1%; Score 716; DB 2; Length 1687;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ADFTFESSTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 60
    |||||||
Db 1120 ADFTFESSTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 1179
Qy 61 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
    |||||||
Db 1180 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 1239
Qy 121 GGARFGLSTENGA 134
    |||||||
Db 1240 GGARFGLSTENGA 1253

RESULT 3
ID Q51816 PRELIMINARY: PRT: 1704 AA.
AC Q51816;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Arg-gingipain-1 proteinase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriis; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA MEHLIN-95138080; PubMed-7636351;
RA Pavloff N., Potempa J., Pike R.N., Prochazka V., Klefer M.C.,
RA Travis J., Barr P.J.;
RT "Molecular cloning and structural characterization of the Arg-
RT gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
RT proteinase-adhesin polypeptide."

```

```

RL J. Biol. Chem. 270:1007-1010(1995).
DR EMBL: U15282; AAA69539.1;
DR HSSP: P95493; ICVR.
DR MEROPS: C25.001;
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ CHAIN 228
FT MATURE 50-KDA CYSTEINE PROTEINASE
FT GINGIPAIN.
SO SEQUENCE 1704 AA: 185436 MW: 6A34B40131C2A676 CRC64:

Query Match
Best Local Similarity 98.1%; Score 716; DB 2; Length 1704;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ADFTFESSTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 60
    |||||||
Db 1137 ADFTFESSTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 1196
Qy 61 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
    |||||||
Db 1197 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 1256
Qy 121 GGARFGLSTENGA 134
    |||||||
Db 1257 GGARFGLSTENGA 1270

RESULT 4
ID P72196 PRELIMINARY: PRT: 1097 AA.
AC P72196;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Tumb-linked adhesin precursor.
GN TLA.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriis; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA MEHLIN-97386416; PubMed-9244265;
RA Aduse-Opoku J., Stanley J.M., Young K.A., Muir J., Rangarajan M.,
RA Curtis M.A.;
RT "The tla gene of Porphyromonas gingivalis W50: a homologue of the
RT arginine-specific protease precursor (Pipri) which shares sequence
RT similarity to 1208-1778."
BL J. Bacteriol. 178:1788-1797(1997).
DR EMBL: Y07618; GA548897.1;
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM SIGNAL.
SQ SEQUENCE 1097 AA: 118731 MW: 73BBB13379421F8B9 CRC64:

Query Match
Best Local Similarity 97.8%; Score 715; DB 2; Length 1097;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ADFTFESSTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 60
    |||||||
Db 522 ADFTFESSTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 581
Qy 61 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
    |||||||
Db 582 LMPDNYLISKDVTGATKRYKYAVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 641
Qy 121 GGARFGLSTENGA 134

```

Db 642 GGARFGLSTEANGA 655

RESULT 5

ID 052050 PRELIMINARY: PRT: 1732 AA.

AC 052050: 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine specific cysteine protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria: CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RP STRAIN=W83;
 RC MEDLINE=98298016; PubMed=9632563;
 RA Lewis J.P., Macrina F.L.,
 RT -15195, an insertion sequence-like element associated with protease
 genes in Porphyromonas gingivalis.";
 RL Infect. Immun. 66:3035-3042(1998).
 DR EMBL: AF017059; AAC26523.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Protease.
 SQ SEQUENCE 1732 AA: 187931 MW: 8233746305CB5EA5 CRC64;

Query Match 97.9%; Score 715; DB 2; Length 1732;
 Best Local Similarity 97.8%; Pred. No. 2,4e-54;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFEESTHGEPAEMWTTIDADGEGMCLSSGQDMLTAAHGTNNVVSFSWNGMA 60
 DB 1157 ADFTETFEESTHGEPAEMWTTIDADGEGMCLSSGQDMLTAAHGTNNVVSFSWNGMA 1216
 OY 61 LNPBNYILSKDVTGATKVKYKYYVPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
 DB 1217 LNPBNYILSKDVTGATKVKYKYYAVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 1276
 OY 121 GGARFGLSTEANGA 134
 DB 1277 GGARFGLSTEANGA 1290

RESULT 6

ID 051817 PRELIMINARY: PRT: 1732 AA.

AC 051817: 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Porphyrypain.
 GN prtp.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria: CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RP STRAIN=W83;
 RC MEDLINE=96213011; PubMed=8631659;
 RA Barkocy-Gallagher G.A., Han N., Patel J.M., Whitlock J.,
 RT "Analysis of the prtp gene encoding porphyrypain, a cysteine proteinase
 of Porphyromonas gingivalis.";
 RL J. Bacteriol. 178:2734-2741(1996).
 DR EMBL: U42210; AAB06565.1; -.

DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA: 187875 MW: 654271DBEF7BCA64 CRC64;

Query Match 97.9%; Score 715; DB 2; Length 1732;
 Best Local Similarity 97.8%; Pred. No. 2,4e-54;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFEESTHGEPAEMWTTIDADGEGMCLSSGQDMLTAAHGTNNVVSFSWNGMA 60
 DB 1157 ADFTETFEESTHGEPAEMWTTIDADGEGMCLSSGQDMLTAAHGTNNVVSFSWNGMA 1216
 OY 61 LNPBNYILSKDVTGATKVKYKYYVPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
 DB 1217 LNPBNYILSKDVTGATKVKYKYYAVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 1276
 OY 121 GGARFGLSTEANGA 134
 DB 1277 GGARFGLSTEANGA 1290

RESULT 7

ID 007442 PRELIMINARY: PRT: 1732 AA.

AC 007442: 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine-specific cysteine proteinase.
 GN prtk.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria: CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RP STRAIN=W50;
 RC MEDLINE=99235907; PubMed=10219167;
 RA Staksel N., Cleal S.M., Bhogal P.S., Reynolds E.C.;
 RT "Characterization of a porphyromonas gingivalis gene prtk that encodes
 a lysine-specific cysteine proteinase and three sequence-related
 RT adhesins.";
 RL Oral Microbiol. Immunol. 14:92-97(1999).
 DR EMBL: U75366; AAB60809.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA: 187914 MW: 45D5B9137791703 CRC64;

Query Match 97.9%; Score 715; DB 2; Length 1732;
 Best Local Similarity 97.8%; Pred. No. 2,4e-54;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFEESTHGEPAEMWTTIDADGEGMCLSSGQDMLTAAHGTNNVVSFSWNGMA 60
 DB 1157 ADFTETFEESTHGEPAEMWTTIDADGEGMCLSSGQDMLTAAHGTNNVVSFSWNGMA 1216
 OY 61 LNPBNYILSKDVTGATKVKYKYYVPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
 DB 1217 LNPBNYILSKDVTGATKVKYKYYAVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 1276
 OY 121 GGARFGLSTEANGA 134
 DB 1277 GGARFGLSTEANGA 1290

RESULT 8

ID 051839

```

ID 051839 PRELIMINARY; PRT: 1706 AA.
AC 051839: 051840:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
GN PRTR.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirsztbaum L., Sotgiropoulos C., Jackson C., Cleal S., Slakeski N.,
Reynolds E.C.;
RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
gingivalis W50 encoding a 132 kDa protein that contains an arginine-
specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Biochem. Biophys. Res. Commun. 207:424-431(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RX MEDLINE=96311339; PubMed=8713096;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes
an arginine-specific thiol proteinase and multiple adhesins.";
RL Biochem. Biophys. Res. Commun. 224:605-610(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RA Reynolds E.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RA Slakeski N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR HSSP: P95493; ICVR.
DR InterPro: IPR001769; Peptidase_C25.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT SIGNAL.
FT CHAIN.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 1706 AA: 185626 MW; E8BDP07C913884 CRC64;

Query Match 97.8%; Score 714; DB 2; Length 1706;
Best Local Similarity 97.8%; Pred. No. 2.8e-54;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGTNNVSSFSWNGMA 60
DB 1139 ADFTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGTNNVSSFSWNGMA 1198
OY 61 LNPNDNYLSKDYATGATKVKYKYYPVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1199 LNPNDNYLSKDYATGATKVKYKYYPVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 1258
OY 121 GGARFGLSTEDANGA 134
DB 1259 GGARFGLSTEDANGA 1272

```

```

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
RT "Molecular cloning and characterization of Porphyromonas gingivalis
lys-gingipain.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U54691; AAA99810.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1723 AA: 186831 MW; 4508A7E50197CEBD CRC64;

```

```

Query Match 97.4%; Score 711; DB 2; Length 1723;
Best Local Similarity 97.0%; Pred. No. 5.3e-54;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGTNNVSSFSWNGMA 60
DB 1156 ADFTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGTNNVSSFSWNGMA 1215
OY 61 LNPNDNYLSKDYATGATKVKYKYYPVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1216 LNPNDNYLSKDYATGATKVKYKYYPVNDGFGDHYAVMISKTGTNAGDFTVVEETPNCINK 1275
OY 121 GGARFGLSTEDANGA 134
DB 1276 GGARFGLSTEDANGA 1289

```

```

RESULT 10
P66967 PRELIMINARY; PRT: 1358 AA.
AC P66967;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hemagglutinin.
GN HAGD.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA Han N., Lepine G., Whitlock J., Wojciechowski L., Progniske-Fox A.;
RT "Cloning, sequencing and characterization of hagD, a member of the
HAREP multigene family in Porphyromonas gingivalis.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U68468; AAA49691.1; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1358 AA: 147102 MW; 47FCA0B25B06DEBD CRC64;

Query Match 97.1%; Score 709; DB 2; Length 1358;
Best Local Similarity 97.0%; Pred. No. 5.9e-54;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADFTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGTNNVSSFSWNGMA 60
DB 791 ADFTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGTNNVSSFSWNGMA 850

```



```

OY 61 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKTGNAGDFTVVEETPNCINK 120
DB 851 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKTGNAGDFTVVEETPNCINK 910
OY 121 GGAREGLSTEANGA 134
DB 911 GGAREGLSTEANGA 924

RESULT 11
ID P72194 PRELIMINARY; PRT: 1723 AA.
AC P72194:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lys-gingipalain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97044756; PubMed=8889827;
RC STRAIN=381;
RA Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT Cloning and sequencing of the gene encoding a novel lysine-specific
RT cysteine proteinase (Lys-gingipalain) in Porphyromonas gingivalis:
RT structural relationship with the arginine-specific cysteine proteinase
RT (Arg-gingipalain).
RL J. Biochem. 120:398-406(1996).
DR MEROPS: C25.007; BAA11870.1; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25. 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1723 AA; 187261 MW; 5628963D51493EB CRC64;

Query Match 97.1%; Score 709; DB 2; Length 1723;
Best Local Similarity 97.0%; Pred. No. 7.9e-54;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADFTETFESSHGGEAPAEWTTIDADGCGWCLSSGQDLWLTAHGTVNVSSFSMNGMA 60
DB 1156 ADFTETFESSHGGEAPAEWTTIDADGCGWCLSSGQDLWLTAHGTVNVSSFSMNGMA 1215
OY 61 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKTGNAGDFTVVEETPNCINK 120
DB 1216 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKTGNAGDFTVVEETPNCINK 1275
OY 121 GGAREGLSTEANGA 134
DB 1276 GGAREGLSTEANGA 1289

RESULT 12
ID 051838 PRELIMINARY; PRT: 1706 AA.
AC 051838:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Protease precursor.
GN PRPR1.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
RN 11
RP SEQUENCE FROM N.A.

```

```

BC STRAIN-W50;
RA MEDLINE=96071894; PubMed=7591131;
RX Aduse-Opoku J., Muir J., Stanley J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease
RT antigen (PrPr1) of Porphyromonas gingivalis W50."
RL Infect. Immun. 63:4744-4754(1995).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Rangarajan M., Aduse-Opoku J., Stanley J.M., Young K.A., Curtis M.A.;
RT "The prp1 and the prp2 arginine-specific protease genes of
RT Porphyromonas gingivalis W50 produce five biochemically distinct
RT enzymes."
RL Mol. Microbiol. 23:0-0(1997).
DR EMBL: X82680; CA57997.1; -.
DR HSSP: P95493; ICVR.
DR MEROPS: C25.001; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25. 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL. 1 23 POTENTIAL.
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E36DCD87EDA8CDD CRC64;

Query Match 96.7%; Score 706; DB 2; Length 1706;
Best Local Similarity 97.0%; Pred. No. 1.4e-53;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADFTETFESSHGGEAPAEWTTIDADGCGWCLSSGQDLWLTAHGTVNVSSFSMNGMA 60
DB 1139 ADFTETFESSHGGEAPAEWTTIDADGCGWCLSSGQDLWLTAHGTVNVSSFSMNGMA 1198
OY 61 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKTGNAGDFTVVEETPNCINK 120
DB 1199 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKTGNAGDFTVVEETPNCINK 1258
OY 121 GGAREGLSTEANGA 134
DB 1259 GGAREGLSTEANGA 1272

RESULT 13
ID 09FAJ0 PRELIMINARY; PRT: 925 AA.
AC 09FAJ0:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE putative outer membrane protein PG57.
DE Porphyromonas gingivalis (Bacteroides gingivalis).
OS Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margets M.,
RA Hocking D., Webb E.;
RT "P. gingivalis polypeptides and nucleic acids."
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Ross B.C., Czajkowski L., Hocking D., Margets M., Webb E., Rothel L.,
RA Patterson M., Agius C., Camuglia S., Reynolds E., Barr I.G.;
RT "Identification of vaccine candidates from genomic analysis of
RT Porphyromonas gingivalis."
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY007522; AAG24228.1; -.
DR InterPro: IPR003961; FN_III.

```

DR SMART: SM00060; FN3; 3;
SQ SEQUENCE 925 AA; 103632 MW; 5FF2198D6914DAE8 CRC64;

Query Match 28.1%; Score 205; DB 2; Length 925;
Best Local Similarity 37.7%; Pred. No. 1e-09;

Matches 53; Conservative 14; Mismatches 43; Indels 34; Gaps 7;

QY 1 ADETFEFSST-----HGEAPAEWTTIDAGDGCGLCLSSGQ---- 38
DB 288 AAYDESIESTSVCGCTLMHYATDAILYENFENGVPVNGMVLVIDADGCGFSW----CHYLNA 342
QY 39 LDMTLTAGGTVNVSSFSW-NGMA-LNPNYLSKQVGTATKYKYYPVND-GFPGDHVAV 95
DB 343 YDAEPGNGGCHSLASVYGPVTPPNYLTTPKVEGAKRVKYSTODANMAAEHVAV 402
QY 96 MISKGTNAGDFTVFEET---PNG 117
DB 403 MASTGTAVGDFVILFEETMTAKPTG 428

RESULT 14

Q9KIB3 PRELIMINARY; PRT; 312 AA.

AC 09KIB3:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical outer membrane protein Pg27.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
CX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M50;
RA Ross B., Barr I., Putterson M., Agius C., Rothe L., Margelits M.,
RT Hocking D., Webb E.;
RT "P. gingivalis polypeptides and nucleic acids";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M50;
RA Ross B.C.;
RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF237555; AAF81413.1; -;
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 2.
SQ SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25F5 CRC64;

Query Match 24.7%; Score 180; DB 2; Length 312;

Best Local Similarity 42.0%; Pred. No. 4.2e-08;
Matches 50; Conservative 16; Mismatches 35; Indels 18; Gaps 7;

QY 4 TETFESSTHGEAPAEWTTIDAGDGCGLCLSSGGLDMLTAGGTVNVSSFSWNG--MAL 61
DB 5 SESFESG---IPAIWKITIDADDGYNMHLTN----FTGQSGD-CVSSASVIGVGAL 54
QY 62 NPDNYLSKQVGTAT---KVKYYPVND-GFPGDHVAVMISKGTNACDF-TVFEET 114
DB 55 TPDNYLTPELKLPTALVEIITWCTQDLTAPSEHVAVYSSSTGNNAADFVNLFEET 113

RESULT 15

Q9JZ15 PRELIMINARY; PRT; 218 AA.

AC 09JZ15:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Isomerase, putative.
GN NMB1338.
OS Neisseria meningitidis (serogroup B).

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; Pubmed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J.C., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Dougherty B.A.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Hickey E.K.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouli H., Qin H., Yatchewyan J.,
RA Gill J., Scarlato V., Maignani V., Piza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58";
RL Science 287:1809-1815(2000).
DR EMBL: AE002482; AAF41713.1; -;
DR TIGR: NMB1338; -;
DR InterPro: IPR002529; FAA_hydrolase.
DR Pfam: PF01557; FAA_hydrolase; 1.
KW Complete proteome.
SQ SEQUENCE 218 AA; 23058 MW; BA531FCD9C010BD9 CRC64;

Query Match 12.4%; Score 90.5; DB 16; Length 218;
Best Local Similarity 26.6%; Pred. No. 2.1;
Matches 38; Conservative 14; Mismatches 48; Indels 43; Gaps 6;

QY 23 DADDGEGC-----WCLSSGGLDMLTAGH--GTVNVSSFSWNGM 59
DB 77 DSDGTGECKDILGCVAGVGLDLTARDIQCLKELKGLPWLKAKGFRHSACVSDFAAAGR 136
QY 60 ALNPDNYLSKQVGTATKYK---YYPVNDGFPGDHVAVMISKY---GTNAGDFTV 110
DB 137 IGNEKVFSLKQNGVLLKQRGDTGLMIVPIRE-----ILHKLADYGLGKGD--LV 185
QY 111 FEETPNKNGKGFARGLSTEANG 133
DB 186 FTGTPSGVGAIGAGDNLALELDQ 208

Search completed: June 26, 2003, 00:53:26
Job time : 41.9882 secs

Thu Jun 20 11:58:19 2003

us-09-980-370-8.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:09 ; Search time 3.08824 Seconds

(without alignments)
200.076 Million cell updates/sec

Title: US-09-980-370-8

Perfect score: 108

Sequence: 1 ALMPDNYLSKDYTGATKVKY 21

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2.6/prodata/1/1aa/5A.COMB.pep:*
3: /cgn2.6/prodata/1/1aa/5B.COMB.pep:*
4: /cgn2.6/prodata/1/1aa/6A.COMB.pep:*
5: /cgn2.6/prodata/1/1aa/6B.COMB.pep:*
6: /cgn2.6/prodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	108	100.0	439 2 US-08-570-311-22	Sequence 22, Appl
2	108	100.0	450 2 US-08-570-311-16	Sequence 16, Appl
3	108	100.0	456 2 US-08-570-311-20	Sequence 18, Appl
4	108	100.0	456 2 US-08-570-311-20	Sequence 20, Appl
5	108	100.0	1087 2 US-08-570-311-8	Sequence 8, Appl
6	108	100.0	1087 2 US-08-570-311-8	Sequence 8, Appl
7	108	100.0	1358 2 US-08-570-311-27	Sequence 27, Appl
8	108	100.0	1687 2 US-08-570-311-29	Sequence 29, Appl
9	108	100.0	1704 3 US-08-336-308A-10	Sequence 10, Appl
10	108	100.0	1704 3 US-08-490-931-10	Sequence 10, Appl
11	108	100.0	1732 2 US-08-570-311-10	Sequence 10, Appl
12	108	100.0	1732 2 US-08-353-485-10	Sequence 10, Appl
13	108	100.0	2628 2 US-08-570-311-14	Sequence 14, Appl
14	108	100.0	2628 2 US-08-570-311-14	Sequence 14, Appl
15	47	43.5	183 4 US-09-117-257-42	Sequence 42, Appl
16	47	43.5	183 4 US-09-489-352-42	Sequence 42, Appl
17	47	43.5	187 2 US-08-589-711-2	Sequence 2, Appl
18	47	43.5	187 4 US-09-117-257-8	Sequence 8, Appl
19	47	43.5	187 4 US-09-117-257-9	Sequence 9, Appl
20	47	43.5	187 4 US-09-117-257-15	Sequence 15, Appl
21	47	43.5	187 4 US-09-117-257-23	Sequence 23, Appl
22	47	43.5	187 4 US-09-117-257-25	Sequence 25, Appl
23	47	43.5	187 4 US-09-117-257-32	Sequence 32, Appl
24	47	43.5	187 4 US-09-221-938-2	Sequence 2, Appl
25	47	43.5	187 4 US-08-945-476-9	Sequence 9, Appl
26	47	43.5	187 4 US-08-945-476-15	Sequence 15, Appl
27	47	43.5	187 4 US-08-945-476-23	Sequence 23, Appl

28	47	43.5	187 4 US-08-945-476-25	Sequence 25, Appl
29	47	43.5	187 4 US-08-489-352-8	Sequence 8, Appl
30	47	43.5	187 4 US-08-489-352-15	Sequence 15, Appl
31	47	43.5	187 4 US-08-489-352-23	Sequence 23, Appl
32	47	43.5	187 4 US-08-489-352-25	Sequence 25, Appl
33	47	43.5	187 4 US-08-489-352-32	Sequence 32, Appl
34	47	43.5	187 4 US-08-489-352-42	Sequence 42, Appl
35	47	43.5	188 4 US-09-117-257-27	Sequence 27, Appl
36	47	43.5	188 4 US-09-117-257-30	Sequence 30, Appl
37	47	43.5	188 4 US-08-945-476-27	Sequence 27, Appl
38	47	43.5	188 4 US-08-945-476-27	Sequence 27, Appl
39	47	43.5	188 4 US-08-489-352-30	Sequence 30, Appl
40	47	43.5	188 4 US-08-489-352-44	Sequence 44, Appl
41	47	43.5	188 4 US-09-117-257-36	Sequence 36, Appl
42	47	43.5	191 4 US-09-117-257-36	Sequence 36, Appl
43	47	43.5	191 4 US-08-489-352-36	Sequence 36, Appl
44	47	42.6	497 2 US-08-570-311-2	Sequence 2, Appl
45	46	42.6	497 2 US-08-353-485-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-570-311-22
Sequence 22, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguis-Ske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepline, Guyalane
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis genes
TITLE OR INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570.311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UP15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:

Thu Jun 26 11:58:19 2003

us-09-980-370-8.ra1

Page 2

LENGTH: 439 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-22

Query Match 100.0%; Score 108; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 251 ALNPDNYLSKDVGTGATKVKY 271

RESULT 2
US-08-570-311-16

Sequence 16, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naiming

APPLICANT: Lantzi, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESS: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UP15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 372-5800

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 450 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-16

Query Match 100.0%; Score 108; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 245 ALNPDNYLSKDVGTGATKVKY 265

RESULT 3
US-08-570-311-18

Sequence 18, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naiming

APPLICANT: Lantzi, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESS: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UP15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 372-5800

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-18

Query Match 100.0%; Score 108; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 251 ALNPDNYLSKDVGTGATKVKY 271

RESULT 4

US-08-570-311-20

Sequence 20, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepline, Guyalaine

APPLICANT: Han, Naïming

APPLICANT: Lantze, Marilyn

APPLICANT: Patel, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 372-5800

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-20

Query Match

Best Local Similarity 100.0%; Score 108; DB 2; Length 456;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNVLISKDVYGAATKRY 21

DB 251 ALNPDNVLISKDVYGAATKRY 271

RESULT 5

US-08-570-311-8

Sequence 8, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepline, Guyalaine

APPLICANT: Han, Naïming

APPLICANT: Lantze, Marilyn

APPLICANT: Patel, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

APPLICANT: Lantze, Marilyn

APPLICANT: Patel, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 372-5800

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1087 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-8

Query Match

Best Local Similarity 100.0%; Score 108; DB 2; Length 1087;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNVLISKDVYGAATKRY 21

DB 579 ALNPDNVLISKDVYGAATKRY 599

RESULT 6

US-08-353-485-8

Sequence 8, Application US/08353485

Patent No. 5830710

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepline, Guyalaine

APPLICANT: Han, Naïming

APPLICANT: Lantze, Marilyn

APPLICANT: Patel, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Query Match 100.0%; Score 108; DB 2; Length 1087;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNLIISKDVTGATKVKY 21
|||||
DB 579 ALNPDNLIISKDVTGATKVKY 599

RESULT 7
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Progulsk-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyline
APPLICANT: Han, Naiming
APPLICANT: Lantiz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: 08/570,311

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ. ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

Query Match 100.0%; Score 108; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNLIISKDVTGATKVKY 21
|||||
DB 850 ALNPDNLIISKDVTGATKVKY 870

RESULT 8
US-08-570-311-29
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Progulsk-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyline
APPLICANT: Han, Naiming
APPLICANT: Lantiz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: 08/570,311
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whillock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Query Match 100.0%; Score 108; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVYGTATKVKY 21
DB 1179 ALNDPNYLISKDVYGTATKVKY 1199

RESULT 9
US-08-336-308A-10
Sequence 10, Application US/08336308A
Patent No. 6017532
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barry, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Feibert, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 100.0%; Score 108; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVYGTATKVKY 21
DB 1196 ALNDPNYLISKDVYGTATKVKY 1216

RESULT 10
US-08-822-324-6
Sequence 6, Application US/08822324
Patent No. 6129917
GENERAL INFORMATION:
APPLICANT: Potempa, Jan S.
APPLICANT: Travis, James
APPLICANT: Genco, Caroline A.
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,945
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Feibert, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 103-95 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 488-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-324-6

Query Match 100.0%; Score 108; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVYGTATKVKY 21
DB 1196 ALNDPNYLISKDVYGTATKVKY 1216

RESULT 11
US-09-490-931-10
Sequence 10, Application US/09490931
Patent No. 6274718
GENERAL INFORMATION:

APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,931
CLASSIFICATION:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/336,308
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-490-931-10

Query Match 100.0%; Score 108; DB 4; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDYVGATKVKY 21
DB 1196 ALNPDNYLSKDYVGATKVKY 1216

RESULT 12
US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naïming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
CLASSIFICATION: 424
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10

Query Match 100.0%; Score 108; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDYVGATKVKY 21
DB 1216 ALNPDNYLSKDYVGATKVKY 1236

RESULT 13
US-08-353-485-10
Sequence 10, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naïming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485

FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-10

Query Match 100.0%; Score 108; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 1216 ALNPDNYLSKDVGTGATKVKY 1236

RESULT 14
US-08-570-311-14
Sequence 14, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguis-Ske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Leplue, Guyalene
APPLICANT: Han, Nalinng
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Query Match 100.0%; Score 108; DB 2; Length 2628;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 744 ALNPDNYLSKDVGTGATKVKY 764

RESULT 15
US-09-117-257-42
Sequence 42, Application US/09117257
Patent No. 6214355
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
APPLICANT: Guo, Betty
APPLICANT: Hanson, Mark
TITLE OF INVENTION: DBPA AND DBPB COMPOSITIONS AND METHODS OF USE
FILE REFERENCE: 4210.000500
CURRENT APPLICATION NUMBER: US/09/117,257
EARLIER FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: PCT/US96/17081
EARLIER FILING DATE: 1996-10-22
EARLIER APPLICATION NUMBER: 08/589,711
EARLIER FILING DATE: 1996-01-22
EARLIER APPLICATION NUMBER: 08/427,023
EARLIER FILING DATE: 1995-04-24
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 42
LENGTH: 183
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-117-257-42

Query Match 43.5%; Score 47; DB 4; Length 183;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 NYLSKDVGTGATKVK 20
DB 20 NLLISGILGATKVK 34

Search completed: June 26, 2003, 00:53:58
Job time : 3.08824 secs

Thu Jun 26 11:58:19 2003

us-09-980-370-8.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:39 ; Search time 4.44706 Seconds

(without alignments)
510.977 Million cell updates/sec

Title: US-09-980-370-8

Perfect score: 108

Sequence: 1 ALNPNTYLSKDYTGATKVKY 21

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubpaa/PCIT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/PCITUS_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	46.3	580	9	US-09-928-457-37
2	44	40.7	81	10	US-09-529-063-19
3	44	40.7	575	9	US-09-738-626-4263
4	43	39.8	381	10	US-09-815-242-10962
5	43	39.8	480	9	US-09-464-099A-49
6	43	39.8	480	10	US-09-861-696-49
7	43	39.8	1588	9	US-10-214-766-34
8	42.5	39.4	369	9	US-09-804-291-457
9	42.5	39.4	369	10	US-09-886-055-457
10	42	38.9	2548	10	US-09-881-682A-1
11	41.5	38.4	491	10	US-09-925-299-1177
12	41	38.0	166	9	US-09-925-299-1177
13	41	38.0	166	10	US-09-925-299-1177
14	41	38.0	431	9	US-09-860-670-117
15	41	38.0	773	9	US-09-808-602-60
16	41	38.0	773	9	US-09-800-198-49
17	41	38.0	850	9	US-09-808-602-58
18	41	38.0	850	9	US-09-800-198-47
19	41	38.0	983	9	US-09-808-602-73

20	41	38.0	983	9	US-10-013-136-2	Sequence 2, Appl
21	41	38.0	1572	9	US-10-037-182-20	Sequence 20, Appl
22	41	38.0	1572	9	US-10-037-182-16	Sequence 16, Appl
23	41	38.0	1605	9	US-10-037-182-18	Sequence 18, Appl
24	41	38.0	1607	10	US-09-938-275-10	Sequence 10, Appl
25	41	38.0	1609	9	US-10-037-182-14	Sequence 14, Appl
26	41	38.0	1609	9	US-10-037-182-12	Sequence 12, Appl
27	41	38.0	1609	10	US-09-938-275-11	Sequence 11, Appl
28	40.5	37.5	374	10	US-09-923-9980-27	Sequence 27, Appl
29	40.5	37.5	324	10	US-09-923-9980-21	Sequence 21, Appl
30	40.5	37.5	324	10	US-09-923-9980-49	Sequence 49, Appl
31	40.5	37.5	324	10	US-09-923-9980-12	Sequence 12, Appl
32	40	37.0	1765	9	US-10-037-182-6	Sequence 6, Appl
33	40	37.0	1765	9	US-10-037-182-10	Sequence 10, Appl
34	40	37.0	1786	9	US-10-037-182-6	Sequence 6, Appl
35	40	37.0	1786	9	US-10-037-182-10	Sequence 10, Appl
36	40	37.0	1786	10	US-09-938-275-6	Sequence 6, Appl
37	40	37.0	1786	10	US-09-938-275-7	Sequence 7, Appl
38	40	37.0	1786	10	US-10-106-698-6717	Sequence 6717, Ap
39	39.5	36.6	158	9	US-10-106-698-6717	Sequence 6717, Ap
40	39.5	36.6	324	10	US-09-815-242-11345	Sequence 11345, A
41	39.5	36.6	349	9	US-10-270-875-57	Sequence 57, Appl
42	39.5	36.6	349	9	US-10-270-878-57	Sequence 57, Appl
43	39.5	36.6	349	9	US-10-270-786-57	Sequence 57, Appl
44	39.5	36.6	349	9	US-10-270-710-57	Sequence 57, Appl
45	39.5	36.6	349	9	US-10-270-859-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-09-928-457-37
Sequence 37, Application US/09928457
Patent No. US20020164603A1
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA, specific proteins and peptides
TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
NUMBER OF SEQUENCES: 99
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30 (OEB)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/928.457
FILING DATE: 2001-08-14
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/214,759
FILING DATE: 199-12-10
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 acids amin.s
TYPE: acids amin.s
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..580
US-09-928-457-37
Query Match: 46.3% Score 50: DB 9: Length 580:
Best local similarity 45.0% Pred. No. 5.3:
Matches 9: Conservative 5: Mismatches 6: Indels 0: Gaps 0:
Cy 2 LNPNNTYLSKDYTGATKVKY 21
Db 56 LKNNYLSKDYTGATKVKY 75

RESULT 2
 US-09-529-063-19
 Sequence 19, Application US/09529063
 Patent No. US20020102542A1
 GENERAL INFORMATION:
 APPLICANT: FUKUSHIMA, DAICHI
 APPLICANT: SHIBAYAMA, SHIRO
 APPLICANT: TADA, HIDEAKI
 TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
 TITLE OF INVENTION: THE BOTH
 FILE REFERENCE: 058769
 CURRENT APPLICATION NUMBER: US/09/529,063
 CURRENT FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: PCT/JP98/04514
 PRIOR FILING DATE: 1998-10-06
 PRIOR APPLICATION NUMBER: JP 9-274674
 NUMBER OF SEQ ID NOS: 117
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 19
 LENGTH: 81
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-529-063-19

Query Match	40.78;	Score 44;	DB 10;	Length 81;
Best Local Similarity	53.38;	Pred. No. 5;		
Matches	8;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0

```
QY      4 PDNYLISKDVTGATK 18
         |::|||::|||
Db      28 PNSYLSKKIAGLTK 42
```

```

RESULT 3
US-09-738-626-4263
; Sequence 4263, Application US/097386620
; Publication No. US20020197605A1
; GENERAL INFORMATION:

```

```

1  APPLICANT: ANDO, SEIRO
2  APPLICANT: HAYASHI, MIKIRO
3  APPLICANT: OCHIAI, KEIKO
4  APPLICANT: YOKOI, HARUHIKO
5  APPLICANT: TATEISHI, NAOKO
6  APPLICANT: SENOH, AKIHITO
7  APPLICANT: IKEDA, MASATO
8  APPLICANT: OZATI, AKIO
9  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
10 FILE REFERENCE: 249-125
11 CURRENT APPLICATION NUMBER: US/09/738,628
12 CURRENT FILING DATE: 2000-12-18
13 PRIOR APPLICATION NUMBER: JP 99/377484
14 PRIOR FILING DATE: 1999-12-16
15 PRIOR APPLICATION NUMBER: JP 00/159162
16 PRIOR FILING DATE: 2000-04-07
17 PRIOR APPLICATION NUMBER: JP 00/280988
18 PRIOR FILING DATE: 2000-08-03
19 NUMBER OF SEQ ID NOS: 7059
20 SOFTWARE: Patencin ver. 3.0
21 SEQ ID NO 4263
22 LENGTH: 575
23 TYPE: PRT
24 ORKANSIM: Corynebacterium glutamicum
25 US-09/738-626-4263

```

```

Query Match      40.7%  Score 44:  DB 9:  Length 575;
Best Local Similarity 42.9%  Pred. No. 52;
Matches 9:  Conservative 4:  Mismatches 8:  Indels 0:  Gaps 0

```

Db 91 ALNPFVPHLVPEDETGIGRALY 111

```

RESULT 4
US-09-815-242-10962
Sequence 10962, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
Applicant: Haselbeck, Robert
Applicant: Ohlson, Karl L.
Applicant: Zyskind, Judith W.
Applicant: Wall, Daniel
Applicant: Trawick, John D.
Applicant: Carr, Grant J.
Applicant: Yamamoto, Robert T.
Applicant: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA 011A
FILE REFERENCE: Pokryshes
CURRENT FILING DATE: US/09/815,242
CURRENT FILING DATE: 2001-03-21
Prior APPLICATION NUMBER: 60/121,078
Prior FILING DATE: 2000-03-21
Prior APPLICATION NUMBER: 60/206,848
Prior FILING DATE: 2000-05-23
Prior APPLICATION NUMBER: 60/207,727
Prior FILING DATE: 2000-05-26
Prior APPLICATION NUMBER: 60/242,578
Prior FILING DATE: 2000-10-23
Prior APPLICATION NUMBER: 60/253,625
Prior FILING DATE: 2000-11-27
Prior APPLICATION NUMBER: 60/257,931
Prior FILING DATE: 2000-12-22
Prior APPLICATION NUMBER: 60/269,308
Prior FILING DATE: 2001-02-16
NUMBER OF SEQ. ID NOS: 1410
SOFTWARE: PASSED for Windows Version 4.0
SEQ. ID NO 10962
LENGTH: 381
TYPE: PRT
ORGANISM: Haemophilus Influenzae
US-09-815-242-10962

```

Query Match	39.88;	Score 43;	DB 10;	Length 381;
Best Local Similarity	50.08;	Pred. No. 47;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0

```
QY      4 PDNYLISKDVTGAT 17
        ||::|||
Db     104 PDEWIVSSDVYGGT 117
```

RESULT 5
 US-09-464-099A-49
 Sequence 49, Application US/09464099A
 Patent No. US20020168680A1
 GENERAL INFORMATION:
 APPLICANT: Barry, Gerard F.
 APPLICANT: Kishore, Ganesh M.
 APPLICANT: Pedgette, Stephen R.
 APPLICANT: Stallings, William C.
 INVENTION: ELUSTRATED TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHAS
 FILE REFERENCE: US09/094640, MO88:115-2
 CURRENT APPLICATION NUMBER: US09/464,099A
 CURRENT FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: US 09/137,440
 PRIOR FILING DATE: 1998-08-20
 PRIOR APPLICATION NUMBER: US 08/833,485
 PRIOR FILING DATE: 1997-04-07
 PRIOR APPLICATION NUMBER: US 08/306,063
 PRIOR FILING DATE: 1994-09-13
 PRIOR APPLICATION NUMBER: US 07/749,611
 PRIOR FILING DATE: 1991-08-28

PRIOR APPLICATION NUMBER: US 07/576,537
PRIOR FILING DATE: 1990-08-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49
LENGTH: 480
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-464-099A-49

Query Match 39.8%; Score 43; DB 9; Length 480;
Best Local Similarity 43.8%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDVTCAT 17
: 11 1:1 1:1
Db 247 IMPSEYIESDASSAT 262

RESULT 6
US-09-861-696-49
Sequence 49, Application US/09861696
Patent No. US2002007053A1
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
FILE REFERENCE: 11899.0175.CNUS04 MOB:175-4
CURRENT APPLICATION NUMBER: US/09/861,696
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/137,440
PRIOR FILING DATE: 1998-08-20
PRIOR APPLICATION NUMBER: US 08/833,485
PRIOR FILING DATE: 1997-04-07
PRIOR APPLICATION NUMBER: US 08/306,063
PRIOR FILING DATE: 1994-09-13
PRIOR APPLICATION NUMBER: US 07/749,611
PRIOR FILING DATE: 1991-08-28
PRIOR APPLICATION NUMBER: US 07/576,537
PRIOR FILING DATE: 1990-08-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49
LENGTH: 480
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-861-696-49

Query Match 39.8%; Score 43; DB 10; Length 480;
Best Local Similarity 43.8%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDVTCAT 17
: 11 1:1 1:1
Db 247 IMPSEYIESDASSAT 262

RESULT 7
US-10-214-766-34
Sequence 34, Application US/10214766
Publication No. US20030084473A1
GENERAL INFORMATION:
APPLICANT: Gocal, Greg
TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
FILE REFERENCE: CA1138
CURRENT APPLICATION NUMBER: US/10/214,766
CURRENT FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,734
PRIOR FILING DATE: 2001-08-09
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn version 3.2

SEQ ID NO 34
LENGTH: 1588
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-214-766-34

Query Match 39.8%; Score 43; DB 9; Length 1588;
Best Local Similarity 43.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDVTCAT 17
: 11 1:1 1:1
Db 645 IMPSEYIESDASSAT 660

RESULT 8
US-09-804-291-457
Sequence 457, Application US/09804291
Publication No. US20030088059A1
GENERAL INFORMATION:
APPLICANT: ZOZULIA, SERGEY
TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: P 0278005
CURRENT APPLICATION NUMBER: US/09/804,291
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/188,914
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/192,033
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/198,474
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/199,335
PRIOR FILING DATE: 2000-04-24
PRIOR APPLICATION NUMBER: 60/207,702
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/213,849
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/226,534
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: 60/230,732
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 60/266,862
PRIOR FILING DATE: 2001-02-07
NUMBER OF SEQ ID NOS: 529
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 457
LENGTH: 369
TYPE: PRT
ORGANISM: Homo sapiens
US-09-804-291-457

Query Match 39.4%; Score 42.5; DB 9; Length 369;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 2 LNPDMYLSKDVTCAT 18
: 11 1:1 1:1
Db 335 LNPYLSLRNKKDVTCAT 352

RESULT 9
US-09-886-055-457
Sequence 457, Application US/09886055
Patent No. US20020132273A1
GENERAL INFORMATION:
APPLICANT: STRYER, LUBERT
APPLICANT: ZOZULIA, SERGEY
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
FILE REFERENCE: 078003-0277150
CURRENT APPLICATION NUMBER: US/09/886,055
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 60/213,812

```

: PRIOR FILING DATE: 2000-06-22
: NUMBER OF SEQ ID NOS: 522
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 457
: LENGTH: 369
: TYPE: PRF
: ORGANISM: Homo sapiens
US-09-886-055-457

Query Match
Best Local Similarity 39.4%; Score 42.5; DB 10; Length 369;
Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 2 LNPDMY-LISKDYTGATK 18
DB 335 LNPDIYSLRNKDYTGALK 352

RESULT 10
US-09-851-682A-1
: Sequence 1, Application US/09851682A
: Patent No. US20020091248A1
: GENERAL INFORMATION:
: APPLICANT: Adams, Arwen E.
: APPLICANT: Chiu, Choi Ying
: APPLICANT: Duhl, David
: APPLICANT: Gorman, Susan W.
: APPLICANT: Leng, Song
: APPLICANT: Sheffield, Val
: TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
: TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
: TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
: FILE REFERENCE: 200130.442
: CURRENT APPLICATION NUMBER: US/09/851,682A
: CURRENT FILING DATE: 2001-05-08
: PRIOR APPLICATION NUMBER: US/09/172,422
: PRIOR FILING DATE: 1998-10-14
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2548
: TYPE: PRF
: ORGANISM: Homo sapien
US-09-851-682A-1

Query Match
Best Local Similarity 38.9%; Score 42; DB 10; Length 2548;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LNPDMY-LISK 11
DB 989 LNPDMYOVGR 998

RESULT 11
US-09-881-752A-128
: Sequence 128, Application US/09881752A
: Patent No. US20020115078A1
: GENERAL INFORMATION:
: APPLICANT: Kleathous, Harold
: APPLICANT: Al-Garawi, Amal
: APPLICANT: Miller, Charles
: APPLICANT: Tomb, Jean-Francois
: APPLICANT: Oomen, Raymond P.
: TITLE OF INVENTION: Identification of Polynucleotides
: TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
: FILE REFERENCE: 06132/041002
: CURRENT APPLICATION NUMBER: US/09/881,752A
: CURRENT FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: US 08/833,457
: PRIOR FILING DATE: 1997-04-01
```

```

: NUMBER OF SEQ ID NOS: 370
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 128
: LENGTH: 491
: TYPE: PRF
: ORGANISM: Helicobacter pylori
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: 84
: OTHER INFORMATION: Xaa = Any Amino Acid
US-09-881-752A-128

Query Match
Best Local Similarity 38.4%; Score 41.5; DB 10; Length 491;
Matches 11; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

OY 2 LNPDMY--LISKDYTGATK 18
DB 173 LNPDMFLGLTKLDYDRAKK 192

RESULT 12
US-09-925-299-1177
: Sequence 1177, Application US/09925299
: Publication No. US20030040617A9
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102
: CURRENT APPLICATION NUMBER: US/09/925,299
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1556
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 1177
: LENGTH: 166
: TYPE: PRF
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (157)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (158)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (163)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1177

Query Match
Best Local Similarity 38.0%; Score 41; DB 9; Length 166;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ALNPDMYLSKDY 13
DB 141 ALNOKNFLISDEL 153

RESULT 13
US-09-925-299-1177
: Sequence 1177, Application US/09925299
: Patent No. US20020055627A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102
: CURRENT APPLICATION NUMBER: US/09/925,299
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
```

```

? PRIOR FILING DATE: 2000-03-08
? PRIOR APPLICATION NUMBER: 60/124,270
? PRIOR FILING DATE: 1999-03-12
? NUMBER OF SEQ ID NOS: 1556
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1177
? LENGTH: 166
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (157)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (158)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (163)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? US-09-925-299-1177

```

Query Match	38.08;	Score 41;	DB 10;	Length 166;
Best Local Similarity	61.58;	Pred. No. 37;		
Matches	8;	Conservative	3;	Mismatches 2;
			Indels	0;
			Gaps	0;

```

OY      1 ALNPBNYLISKDV 13
        111 1:111:1:
Db      141 ALNOKNELISED, 153

```

```

RESULT 14
US-09-860-670-117
: Sequence 117, Application US/09860670
: Patent No. US20020165137A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PA127P1
: CURRENT APPLICATION NUMBER: US/09/860,670
: PRIORITY FILING DATE: 2001-05-21
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 289
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 117
: LENGTH: 431
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (205)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-860-670-117

```

Query Match	38.0%	Score 41	DB 9	Length 431
Best Local Similarity	50.0%	Pred. No. 1.2e+02		
Matches	8	Conservative	2	Mismatches 6
				Indels 0
				Gaps 0

Qy 6 NYLISKDVTGATKKVKY 21
:|:| | |||
Db 388 HYIISASSDGTVKVKY 403

RESULT 15
 US-09-808-602-60
 : Sequence 60, Application US/09808602
 : Patent No. US20020151151A1
 : GENERAL INFORMATION:
 :
 : APPLICANT: Vernet, Corinne A
 : APPLICANT: Bernades, Elma
 : APPLICANT: SHIMKETS, Richard A
 : APPLICANT: Herriman, John L
 : APPLICANT: Wejunder, Kunud
 : APPLICANT: Mishra, Vishnu

```

: APPLICANT: Mezes, Peter S
: APPLICANT: Macdougall, John
: TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 15966-697 CIP
: CURRENT APPLICATION NUMBER: US/09/808,602
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 09/800,198
: PRIOR FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,596
: PRIOR FILING DATE: 2000-03-03
: NUMBER OF SEQ ID NOS: 114
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 60
: :
: LENGTH: 773
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-808-602-60

```

Query Match	38.08;	Score 41;	DB 9;	Length 773;
Best Local Similarity	45.58;	Pred. No. 2.3e+02;		
Matches 10; Conservative	6;	Mismatches 4;	Indels 2;	Gaps 2

```
QY      2 LNPD-NYLIS-KDVTGATKKVY 21
        ::||::|||::|||
Db      625 VSPDGHYLVISINDVKGLVRVQY 646
```

Search completed: June 26, 2003, 00:54:45
Job time : 4.44706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:46:24 ; Search time 3.45882 Seconds

(Without alignments)
583.673 Million cell updates/sec

Title: US-09-980-370-8

Sequence: 1 ALNPDNYLISKDVTGATKVKY 21

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
2	108	100.0	1704	2	gingipain R (EC 3.4.22.37)
3	108	100.0	1732	2	lysine-specific cy
4	108	100.0	2628	2	hemagglutinin A -
5	35	50.9	334	2	hypothetical prote
6	30	46.3	580	2	hemolysin activati
7	30	46.3	580	2	hemolysin activati
8	49	45.4	265	2	probable periplasm
9	46	42.6	326	2	luteolin gamma chal
10	45	41.7	358	2	climanyl alcohol pr
11	44	40.7	191	2	H70209
12	44	40.7	315	2	B96533
13	44	40.7	390	2	B84553
14	44	40.7	633	2	B84553
15	44	40.7	639	2	B84553
16	44	40.7	639	2	B84553
17	44	40.7	639	2	B84553
18	44	40.7	639	2	B84553
19	44	40.7	639	2	B84553
20	44	40.7	639	2	B84553
21	44	40.7	639	2	B84553
22	44	40.7	639	2	B84553
23	44	40.7	639	2	B84553
24	44	40.7	639	2	B84553
25	44	40.7	639	2	B84553
26	44	40.7	639	2	B84553
27	44	40.7	639	2	B84553
28	44	40.7	639	2	B84553
29	44	40.7	639	2	B84553

30	42.5	39.4	1965	2	T33216	hypothetical prote
31	42	38.9	169	2	C81192	hypothetical prote
32	42	38.9	232	2	P97914	conserved hypothet
33	42	38.9	310	2	B97777	chlorodioxin-disulf
34	42	38.9	318	2	D86325	hypothetical prote
35	42	38.9	383	2	S38904	hypothetical prote
36	42	38.9	528	2	H70163	phenylalanine-LNA
37	42	38.9	561	2	T34368	hypothetical prote
38	42	38.9	653	2	T01274	hypothetical prote
39	42	38.9	983	1	VCLJVS	env polypeptide pr
40	42	38.9	983	1	E45390	hemolysin A toxin
41	42	38.9	998	2	T00227	myosin IXA [import
42	42	38.9	1777	2	T34369	cytochrome P450-11
43	42	38.9	2548	2	E59435	hypothetical prote
44	41.5	38.4	512	2	T48462	hypothetical prote
45	41.5	38.4	704	2	D97723	hypothetical prote

ALIGNMENTS

RESULT 1

S49763 gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (Fragment)

C:Species: Porphyromonas gingivalis

C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997

C:Accession: S49763

R:Aduse-Opoku, J.; Mull, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A. submitted to the EMBL Data Library, November 1994

A:Description: Cloning, sequence analysis and expression in Escherichia coli of prpRI

A:Reference number: S49763

A:Accession: S49763

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1526 <ADU>

A:Cross-references: EMBL:X82680

A:Genetics:

A:Gene: prpRI

C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0% Score 108; DB 2; Length 1526;

Best Local Similarity 100.0%; Pred. No. 8.8e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTGATKVKY 21

DB 1198 ALNPDNYLISKDVTGATKVKY 1218

RESULT 2

A53426 gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis

N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP

C:Species: Porphyromonas gingivalis

C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999

C:Accession: A53426; D53113

R:Payloft, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr

J. Biol. Chem. 270, 1007-1010, 1995

A:Title: Molecular cloning and structural characterization of the Arg-gingipain prote

A:Reference number: A53426; MIMD:95138080; PMID:7636351

A:Accession: A53426

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1704 <PAV>

A:Cross-references: GDB:U5282; NID:9557067; PUDN:AAA69539.1; PID:9557068

R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.

J. Biol. Chem. 269, 406-411, 1994

A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso

A:Reference number: A53113; MIMD:94103245; PMID:8276827

A:Accession: D53113

A:Status: preliminary

A:Molecule type: protein

A:Residues: 228-249 <PIK>

Query Match 46.3%; Score 50; DB 2; Length 580;
Best Local Similarity 45.0%; Pred. No. 6.5;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 2 LMPDNYLISKDVTGATKVKY 21
DB 56 LKNNYVLSEDETPCTRVNT 75

RESULT 7

AB1989 Probable periplasmic protein NMA0687 [Imported] - Neisseria meningitidis (strain Z2491)

C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: AB1989
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1773; MUID:2022356; PMID:10761919
A:Accession: AB1989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-580 <PAR>
A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83973.1; PID:g737941
A:Experimental source: serogroup A, strain Z2491
C:Genetics: NMA0686; NMA0687

Query Match 46.3%; Score 50; DB 2; Length 580;
Best Local Similarity 45.0%; Pred. No. 6.5;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 2 LMPDNYLISKDVTGATKVKY 21
DB 56 LKNNYVLSEDETPCTRVNT 75

RESULT 8

S40209 tubulin gamma chain - fungus (Cochliobolus heterostrophus)

C:Species: Cochliobolus heterostrophus, Bipolaris maydis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
R:Parkinson, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlin, M.H.
A:Submitted to the EMBL Data Library, August 1993
A:Description: Phylogenetic analyses using the gamma tubulin gene.
A:Reference number: S40209
A:Accession: S40209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <PAR>
A:Cross-references: EMBL:X74455; NID:g437988; PIDN:CAA52464.1; PID:g437989
C:Genetics:
A:Introns: 136/3
C:Superfamily: tubulin

Query Match 45.4%; Score 49; DB 2; Length 265;
Best Local Similarity 64.3%; Pred. No. 4;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 NPDNYLISKDVTGTA 16
DB 22 NPDNYLISKDVTGTA 35

RESULT 9

B46108 Outer capsid protein VP7 - human rotavirus (strain 1321)

C:Species: human rotavirus
C:Date: 07-Apr-1994 #sequence_revision 25-Apr-1997 #text_change 20-Sep-1999
C:Accession: B46108

R:Das, M.; Dunn, S.J.; Woode, G.N.; Greenberg, H.B.; Rao, C.D.

Virology 194, 374-379, 1993

A:Title: Both surface proteins (VP4 and VP7) of an asymptomatic neonatal rotavirus st

A:Reference number: A46108; MUID:93242771; PMID:8386881

A:Contents: 1321, serotype 10

A:Accession: B46108

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-326 <DAS>

A:Cross-references: GB:L07658; NID:g310326; PIDN:AAA47308.1; PID:g310327

A:Note: sequence extracted from NCBI backbone (NCBI:130701, NCBI:130693)

A:Note: the sequence in Genbank entry R01VP7B, release 109, (PID:g310325) has the cod

C:Superfamily: rotavirus glycoprotein VP7

Query Match 42.6%; Score 46; DB 2; Length 326;
Best Local Similarity 40.0%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 2 LMPDNYLISKDVTGATKVKY 21
DB 48 VNPQVGINLPITGSDMSY 67

RESULT 10

S31572

cinamyl-alcohol dehydrogenase (EC 1.1.1.165) - alfalfa

C:Species: Medicago sativa (alfalfa)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 15-Oct-1999

C:Accession: S31572

R:van Doorselaere, J.; van Montagu, M.; Inze, D.

A:Submitted to the EMBL Data Library, January 1993

A:Description: Isolation and characterization of poplar and alfalfa cinamyl alcohol

A:Reference number: S31571

A:Accession: S31572

A:Molecule type: mRNA

A:Residues: 1-358 <VAN>

A:Cross-references: EMBL:219573; NID:g19594; PIDN:CAV7625.1; PID:g19595

C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

C:Keywords: alcohol metabolism; metalloprotein; oxidoreductase; zinc

F:33-341/Domain: long-chain alcohol dehydrogenase homology <LADH>

F:48,70,164/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted

Query Match 41.7%; Score 45; DB 2; Length 358;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 LMPDNYLISKDVTG 15
DB 225 LGADNYLVSSDITVG 238

RESULT 11

H70209

decorin binding protein A (dbpa) homolog - Lyme disease spirochete plasmid A/1p54

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: H70209

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Tadhigra, R.; Wh

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu

lman, C.; Garland, S.; Fujil, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venier, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: H70209

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-191 <KLE>

A:Cross-references: GB:AE000790; NID:g2690224; PIDN:ANC66250.1; PID:g2690249; TIGR:BB

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

C:Superfamily: Lyme disease spirochete plasmid probable decorin binding protein A

Query Match 40.7%: Score 44; DB 2; Length 191;
 Best Local Similarity 60.0%: Pred. No. 18;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 NYLISKDVTGATKVK 20
 DB 20 NLISCGTATKIR 34

RESULT 12

B96533
 Hypothetical protein F14J22.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B96533

R:Methodology: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hultzer, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B96533

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <STO>

A:Cross-references: GB:AE005173; NID:g10120425; PIDN:AA613050.1; GSPDB:GN00141

C:Genetics:

A:Gene: F14J22.12

A:Map position: 1

Query Match 40.7%: Score 44; DB 2; Length 315;
 Best Local Similarity 61.5%: Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ALPNPNYLSKDV 13
 DB 34 SLNPQNDVYSKDV 46

RESULT 13

AB1656
 cytochrome beta-lyase homolog lln1787 [imported] - Listeria innocua (strain Clip11262

C:Species: Listeria innocua

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AB1656

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihi, H.

A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Me

Science 294, 849-852, 2001

A:Title: Comparative genomics of Listeria species

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1656

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-390 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97018.1; PID:g16414274; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lln1787

C:Superfamily: O-succinylhomoserine (thiol)-lyase

OY 5 DNTLISKDVTGAT 17
 DB 93 DHFLISKDVTGAT 105

RESULT 14

B84553
 Probable protein kinase/endonuclease [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: B84553

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84553

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-393 <STO>

A:Cross-references: GB:AE002093; NID:g4914373; PIDN:AA622909.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2G17520

A:Map position: 2

Query Match 40.7%: Score 44; DB 2; Length 393;
 Best Local Similarity 56.2%: Pred. No. 39;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 PDNYLSKDVYATK 19
 DB 125 PQNYLSKDVYATK 140

RESULT 15

A86845
 Hypothetical protein yscA [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: A86845

R:Boletín, A.; Wincker, P.; Manger, S.; Tallon, O.; Malarme, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: A86845

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-659 <STO>

A:Cross-references: GB:AE005176; PID:g12724783; PIDN:AAK05859.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yscA

Query Match 40.7%: Score 44; DB 2; Length 639;
 Best Local Similarity 81.8%: Pred. No. 67;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 PDNYLSKDVYAT 14
 DB 452 PDNYLSKDVYAT 462

Search completed: June 26, 2003, 00:52:22
 Job time: 3.45882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:19 ; Search time 2.22353 Seconds

(without alignments)
391.721 Million cell updates/sec

Title: US-09-980-370-8

Perfect score: 108

Sequence: 1 ALNDNDYLSKDVGTATKVKY 21

Scoring table: Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	989	1	PRTN_PORGI
2	108	100.0	2628	1	HAGA_PORGI
3	49	45.4	265	1	TBG_COCH
4	45	41.7	358	1	CADH_MEDSA
5	45	41.7	418	1	LEU2_THETN
6	45	41.7	462	1	TBG2_EUPCR
7	44	40.7	191	1	DBPA_BORBU
8	44	40.7	789	1	CAD5_HUMAN
9	43.5	40.3	313	1	GAS2_HUMAN
10	43	39.8	209	1	KGUA_CLOAB
11	43	39.8	364	1	CARA_LACPL
12	43	39.8	369	1	METB_HAEIN
13	43	39.8	1588	1	AKO1_YEAST
14	43	39.8	2512	1	FOLN_STNDV
15	43	39.8	2514	1	FOLN_STNDV
16	42.5	39.4	311	1	O2T1_HUMAN
17	42.5	39.4	314	1	GAS2_MOUSE
18	42.5	39.4	851	1	MCEL_FOWPV
19	42	38.9	310	1	TRXB_RICCN
20	42	38.9	528	1	SYFA_BORBU
21	42	38.9	588	1	SYD_PASMU
22	42	38.9	982	1	ENV_VILV
23	42	38.9	983	1	ENV_VILV
24	41	38.0	302	1	LGUL_SCHPO
25	41	38.0	417	1	LE22_THEMA
26	41	38.0	618	1	CHAA_CLOBI
27	41	38.0	710	1	PKNG_MYXXA
28	41	38.0	808	1	SECA_MYXPN
29	41	38.0	1607	1	LMG1_MOUSE
30	41	38.0	1607	1	LMG1_MOUSE
31	40.5	37.5	524	1	LCYE_HUMAN
32	40.5	37.5	817	1	PPSA_PYRAF
33	40.5	37.5	819	1	PPSA_PYRAF

34	40.5	37.5	821	1	PPSA_PYRHO
35	40	37.0	278	1	PR28_MYCLE
36	40	37.0	297	1	EX53_MYCPU
37	40	37.0	354	1	WMD2_HERAU
38	40	37.0	446	1	TBG_SCHRO
39	40	37.0	452	1	TBG_CHLRE
40	40	37.0	468	1	TBG_CHLRE
41	40	37.0	532	1	RRPO_BPGA
42	40	37.0	586	1	VGUN_BREV
43	40	37.0	942	1	AMPN_MANSE
44	40	37.0	1786	1	LMB1_HUMAN
45	40	37.0	1786	1	LMB1_MOUSE

ALIGNMENTS

RESULT 1
ID PRTN_PORGI STANDARD: PRT: 989 AA.
AC P46071:1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protease prth (EC 3.4.22.-).
GN PRTN.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_Taxid=837;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95012612; PubMed=7927685;
RX Fletcher H.M., Schenkeln H.A., Macrina F.L.;
RA "Cloning and characterization of a new protease gene (prth) from
RT Porphyromonas gingivalis."
RL Infect. Immun. 62:4279-4286(1994).
RN [2]
RP ERRATUM.
RA Fletcher H.M., Schenkeln H.A., Macrina F.L.;
RL Infect. Immun. 62:5707-5707(1994).
CC -!- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENABLE
CC P. GINGIVALIS TO EVADE COMPLEMENT-MEDIATED KILLING DURING THE
CC IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS
CC AND IS A VIRULENCE FACTOR.
CC -!- SUBCELLULAR LOCATION: IN MEMBRANE VESICLES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: L27483; AA51298.1; -
CC HSSP: P23882; 1PMT.
CC DR MEMOPS: C25_001; -
CC DR InterPro: IPR001769; Peptidase_C25.
CC DR InterPro: IPR002376; formyl_transf.
CC DR Pfam: PF00551; formyl_transf. 1.
CC DR Pfam: PF01364; Peptidase_C25; 1.
CC FT Hydrolyase; Thiol protease; Repeat; Virulence.
CC REPEAT 270 323
CC FT REPEAT 528 581
CC SEQUENCE 989 AA: 110238 MW: FA85FEBA3AC8944C CRC64:
Query Match 100.0%; Score 108; DB 1; Length 989;
Best Local Similarity 100.0%; Prod. No. 2.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLSKDVGTATKVKY 21
 |||
 DB 149 ALNDPNYLSKDVGTATKVKY 169

RESULT 2

HAGA_PORGI STANDARD; PRT: 2628 AA.

AC 051845: 2000 (Rel. 39, Created)
 DT 30-MAR-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin A precursor.
 GN HAGA.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RX MEDLINE=97047672; PubMed=8926061;
 RA Han N., Whitlock J., Proguiske-Fox A.;
 RT "The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381
 RT contains four large, contiguous, direct repeats."
 RL Infect. Immun. 64:4000-4007(1996).
 CC -1- FUNCTION: AGGLUTINATES ERYTHROCYTES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 041807; AB017128.1;
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 6.
 KW Hemagglutinin; Virulence; Hydrolase; Thiol protease; Signal; Repeat.

FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 2628 HEMAGGLUTININ A.
 FT DOMAIN 25 539 PEPTIDASE C25-LIKE 1.
 FT DOMAIN 540 995 PEPTIDASE C25-LIKE 2.
 FT DOMAIN 996 1451 PEPTIDASE C25-LIKE 3.
 FT DOMAIN 1452 1907 PEPTIDASE C25-LIKE 4.
 FT DOMAIN 2074 2628 PEPTIDASE C25-LIKE 5.
 SO SEQUENCE 2628 AA; 283324 MW; 61C4DE32540C99DA CRC64;

Query Match 100.0%; Score 108; DB 1; Length 2628;
 Best Local Similarity 100.0%; Pred. No. 6; Se-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLSKDVGTATKVKY 21
 |||
 DB 744 ALNDPNYLSKDVGTATKVKY 764

RESULT 3
 TBG_COCHIE STANDARD; PRT: 265 AA.

AC P40633;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tubulin gamma chain (gamma tubulin) (Fragment).
 OS Cochliobolus heterostrophus (Drechslera maydis).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; Cochliobolus.
 OX NCBI_TaxID=5016;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C5;
 RA Parkinson C., Luo H., Knight A., Ahlgvist J., Perlman M.H.;
 RL Submitted (Aug-1993) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
 CC GAMMA TUBULIN IS FOUND AT MICROTUBULE ORGANIZING CENTERS (MTOC)
 CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
 CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X74455; CAA52464.1;
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_Fts2.
 DR Pfam: PF00091; tubulin_1.
 DR PROSITE: PS00227; TUBULIN_1.
 KW Microtubules; GTP-binding.
 FT NON_TER 1 1
 FT MP_BIND 77 83 GTP (POTENTIAL).
 FT NON_TER 265 265
 SO SEQUENCE 265 AA; 29567 MW; A5DA0C23E7D62DC6 CRC64;

Query Match 45.4%; Score 49; DB 1; Length 265;
 Best Local Similarity 64.3%; Pred. No. 1; Se-09;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NPNDYLSKDVGTGA 16
 |||
 DB 22 NPNDYLSKDVGTGA 35

RESULT 4
 CADH_MEDSA STANDARD; PRT: 358 AA.

AC P31656;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) (CAD).
 GN CAD2.
 OS Medicago sativa (Alfalfa).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 OX NCBI_TaxID=3879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Apollo;
 RA van Doorselaere J., Baucher M., Feuillet C., Boudet A.M.,
 RA van Montagu M., Inze D.;
 RT "Isolation of cinnamyl alcohol dehydrogenase cDNAs from two important
 RT economic species: alfalfa and poplar. Demonstration of a high homology
 RT of the gene within angiosperms."
 RL Plant Physiol. Biochem. 33:105-109(1995).
 RN [2]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN=cv. Striver; TISSUE=stem.
 RX MEDLINE=20040095; PubMed=1059494;
 RA Brill E.M., Abraham S., Hayes C.M., Jenkins C.L., Watson J.M.;
 RT "Molecular characterization and expression of a wound-inducible cDNA
 RT encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne
 RT (Medicago sativa L.)."
 RL Plant Mol. Biol. 41:279-291(1999).
 CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FINAL STEP IN A BRANCH OF
 CC PHENYLPROPANOID SYNTHESIS SPECIFIC FOR PRODUCTION OF LIGNIN
 CC MONOMERS. IT ACTS ON CINFERYL-, SINAPYL-, 4-COUMARYL- AND
 CC CINNAMYL-ALCOHOL.

```

CC -1- CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde +
CC NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: Lipid synthesis
CC -1- TISSUE SPECIFICITY: MOST ACTIVELY EXPRESSED IN STEM, HYPOCOTYL AND
CC ROOT TISSUE.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 219573; CAI79625.1; -
CC EMBL: AF083332; AAC35845.1; -
CC PIR: S31572; S31572.
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; Adh_zn_family.
CC Pfam: PF00107; adh_zinc.1.
CC PROSITE: PS00059; ADH_ZINC; 1.
CC Oxidoreductase: NADP; zinc; Lipid biosynthesis.
CC METAL 48 48 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 70 70 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 101 101 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 107 107 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 115 115 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).
CC SO SEQUENCE 358 AA; 38948 MW; FBA609408D01BF56 CRC64;

Query Match 41.7%; Score 45; DB 1; Length 358;
Best Local Similarity 57.1%; Pred. No. 9.9;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 LAMPDNYLSKDVGTG 15
DB 225 LGADNYLVSSDTVG 238

RESULT 5
LEU2_THETN
ID LEU2_THETN STANDARD: PRT: 418 AA.
AC QBRDK2:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-Isopropylmalate dehydratase large subunit (EC 4.2.1.33)
DE (Isopropylmalate isomerase) (Alpha-IPM isomerase) (IPM1).
GN LEUC OR TPE0017
OS Theomanoerobacter tengcongensis.
OC Bacteria: Firmicutes: Clostridia: Thermomanoerobacteriales:
OC Thermomanoerobacteriaceae: Thermomanoerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome.";
RT Genome Res. 12:689-700(2002).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-Isopropylmalate = 2-Isopropylmalate +
CC H(2)O.
CC -1- CATALYTIC ACTIVITY: 2-Isopropylmalate + H(2)O = 2-
CC Isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.

```

```

CC -1- SUBUNIT: Heterodimer of Leuc and Leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE012976; AAM23334.1; -
CC InterPro: IPR001030; Aconitase_N.
CC Pfam: PF00330; aconitase.1.
CC PRODOM: PD000511; Aconitase_N.1.
CC PROSITE: PS00450; ACONITASE_1; 1.
CC PROSITE: PS01244; ACONITASE_2; FALSE_NEG.
CC Leucine biosynthesis: Lyase: Iron-sulfur; 4Fe-4S; Complete proteome.
CC METAL 298 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC METAL 358 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC METAL 361 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC SO SEQUENCE 418 AA; 45510 MW; AC3DE0D6E7A97AE CRC64;

Query Match 41.7%; Score 45; DB 1; Length 418;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAMPDNYLSKDVGTGATV 19
DB 60 ALVDPHFVPSKDIKSAEOV 78

RESULT 6
TBG2_EUPCR
ID TBG2_EUPCR STANDARD: PRT: 462 AA.
AC PS4404:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tubulin gamma-2 chain (Gamma-2 tubulin).
DE Eukaryotes crassus.
OS Eukaryotes: Alveolata: Ciliophora: Spirotrichae: Hypotrichia;
OC Euplotida: Euplotidae; Moneuplotes.
OX NCBI_TaxID=5936;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98192540; PubMed=9524221;
RA Tan M., Heckmann K.;
RA "The two gamma-tubulin-encoding genes of the ciliate Euplotes crassus
RA differ in their sequences, codon usage, transcription initiation
RA sites and poly(A) addition sites.";
RL Gene 210:53-60(1998).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
CC GAMMA TUBULIN IS FOUND AT MICROTUBULE ORGANIZING CENTERS (MTOC)
CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X85235; CAI58490.1; -
CC EMBL: Y09551; CAI70742.1; -
CC InterPro: IPR000217; Tubulin.
CC InterPro: IPR003008; Tubulin_ftsz.
CC Pfam: PF00091; tubulin.1.
CC PRINTS: PR01161; TUBULIN.

```


DR PROSITE: P500227: TUNULIN; 1.
 KM Microtubules: GTP-binding, Multigene family.
 FT NP_BIND 148 GTP (POTENTIAL)
 SO SEQUENCE 462 AA: 52117 MW: FF65B755E9AC846 CRC64.
 Query Match 41.7%; Score 45; DB 1; Length 462;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OY 3 NP0NYLSKDVNGA 16
 DB 87 NPEFVSKTMQGA 100
 RESULT 7
 DBPA_BORBU STANDARD; PRT: 191 AA.
 AC 050917: Q9R805: 006876;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Decorin binding protein A precursor.
 GN DBPA OR BBA24.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Plasmid p954.
 OX Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
 RN NCBI_TaxID=139.
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RA MEDLINE=95369901; PubMed=7642279;
 RA Guo B.P., Norris S.J., Rosenberg L.C., Hook M.;
 RT "Adherence of Borrelia burgdorferi to the proteoglycan decorin";
 RL Infect. Immun. 63:3467-3472(1995).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31, 3028, HBNC, LP7, and SH-2-82;
 RA MEDLINE=98065943; PubMed=9784533;
 RA Roberts W.C., Mullikin B.A., Lathigra R., Hanson M.S.;
 RT "Molecular analysis of sequence heterogeneity among genes encoding
 decorin binding proteins A and B of Borrelia burgdorferi sensu lato";
 RL Infect. Immun. 66:5275-5285(1998).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RA MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Castjens S., Huang W.M., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.R., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utechtack T., Wathay L., McDonald L., Atlach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia
 burgdorferi";
 RL Nature 390:580-586(1997).
 CC -1- FUNCTION: Binds to decorin which may mediate the adherence of
 B. burgdorferi to collagen fibers in skin and other tissues.
 CC -1- SIMILARITY: BELONGS TO THE DECORIN-BINDING PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U75866; AAD05353.1;
 CC EMBL: AF069253; AAC70026.1;
 CC EMBL: AF069255; AAC70030.1;
 CC EMBL: AF069257; AAC70034.1;
 CC -----

DR EMBL: AF069286; AAC70064.1;
 DR EMBL: AF069269; AAC70047.1;
 DR EMBL: AF069275; AAC70053.1;
 DR EMBL: AE000790; AAC66250.1;
 DR TIGR: BBA24;
 DR InterPro: IPR003332; Decorin_bind.
 DR Pfam: PF02352; Decorin_bind; 1.
 KW Signal; Plasmid; Complete proteome.
 FT SIGNAL 1 29
 FT CHAIN 30 191
 FT VARIANT 34 34
 FT VARIANT 37 37
 FT VARIANT 41 43
 FT VARIANT 52 52
 FT VARIANT 55 56
 FT VARIANT 128 128
 FT VARIANT 140 140
 FT VARIANT 169 169
 FT VARIANT 172 173
 FT VARIANT 179 191
 SO SEQUENCE 191 AA: 21213 MW: ADIFL9IC774AFEA1 CRC64;
 Query Match 40.7%; Score 44; DB 1; Length 191;
 Best Local Similarity 60.0%; Pred. No. 7.3;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 OY 6 NYLISKDVGTATVK 20
 DB 20 NLLISGLTGATKIR 34
 RESULT 8
 CAD9_HUMAN STANDARD; PRT: 789 AA.
 ID CAD9_HUMAN
 AC G9ULB4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-9 precursor.
 GN CDH9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=10861224;
 RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
 RT "Identification of three human type-II classic cadherins and frequent
 RT heterophilic interactions between different subclasses of type-II
 RT classic cadherins";
 RL Biochem. J. 349:159-167(2000).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS. CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB035302; BAA87416.1;
 CC HSP: P15116; INCT.

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE007681; AAK79684.1; -
 DR InterPro: IPR000619; Guanylate_kin.
 DR Pfam: PF00625; Guanylate_kin.1.
 DR SMART: SM0072; Gukc.1.
 DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
 DR Transferrase; kinase; ATP-binding; Complete proteome.
 FT NP_BIND 12 19 ATP (BY SIMILARITY).
 FT SEQUENCE 209 AA: 23771 MW: 178646B82843163 CRC64:
 SQ

Query Match 39.8%; Score 43; DB 1; Length 209;
 Best Local Similarity 44.4%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 NPDNYLSKDYTGATKVK 20
 Db 93 NGENVLEIDIOGALKVK 110
 1 : : : 1 : 1 1 1 1

RESULT 11
 CARA_LACPL STANDARD: PRT; 364 AA.
 ID CARA_LACPL
 AC P77885;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbamoyl-phosphate synthase, pyrimidine-specific, small chain (PC
 6.3.5.5) (Carbamoyl-phosphate synthetase glutamine chain) (CPS-P).
 DE PYRAA.
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 CX NCBI_TaxID=1590;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-ATCC 8014 / CCM 1904;
 RX MEDLINE=97136688; PubMed=8982065;
 RA Elgaoez A., Abdi A., Hubert J.-C., Kammerer B.;
 RT "Structure and organisation of the pyrimidine biosynthesis pathway
 genes in Lactobacillus plantarum: a PCR strategy for sequencing
 without cloning."
 RT Gene 182:37-43(1996).
 RL [2]
 RN FUNCTION.
 RP STRAIN-ATCC 8014 / CCM 1904;
 RX MEDLINE=20312804; PubMed=10852872;
 RA Nicoloff H., Hubert J.-C., Brinzel F.;
 RT "In Lactobacillus plantarum, carbamoyl phosphate is synthesized by two
 carbamoyl-phosphate synthetases (CPS): carbon dioxide differentiates
 the arginine-repressed from the pyrimidine-regulated CPS."
 RT J. Bacteriol. 182:3416-3422(2000).
 RL -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
 phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
 CC promotes the hydrolysis of glutamine to ammonia, which is used by
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate.
 CC -1- ENZYME REGULATION: Inhibited by pyrimidine.
 CC -1- SIMILARITY: BELONGS TO THE CARA FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z54240; CA91004.1; -
 DR HSSP: P00907; 1A9X.
 DR InterPro: IPR002474; CPSase_sm_chain.
 DR InterPro: IPR001317; CPS_GATase.
 DR InterPro: IPR000991; GATase_1.
 DR Pfam: PF00988; CPSase_sm_chain.1.
 DR Pfam: PF00117; GATase.1.
 DR PRINTS: PR00099; CPSGATASE.
 DR PRINTS: PR00096; GATASE.
 DR PROSITE: PS00442; GATASE_TYPE_1; 1.
 DR Pyrimidine biosynthesis; Ligase; Glutamine amidotransferase.
 FT DOMAIN 1 165
 FT ACT_SITE 166 364
 FT ACT_SITE 244 244
 FT SEQUENCE 364 AA: 39997 MW: 00B3D7ABC8208A42 CRC64:
 SQ

Query Match 39.8%; Score 43; DB 1; Length 364;
 Best Local Similarity 40.0%; Pred. No. 21;
 Matches 10; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

OY 1 ALNPDNYLIS-----KDYTGATKVK 19
 Db 206 ALNPDGYLITNGPGDPKDPVGALEM 230
 11111 : : : 111 11 : :

RESULT 12
 METB_HAEIN STANDARD: PRT; 369 AA.
 ID METB_HAEIN
 AC P44502;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cystathionine gamma-synthase (EC 4.2.99.9) (CGS) (O-succinylhomoserine
 (Thiol)-lyase).
 DE METB OR H10086.
 GN Haemophilus influenzae.
 OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 CX NCBI_TaxID=727;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-Rd / RW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA McKenney K., Sutton G., Fitch J.W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd."
 RT Science 269:496-512(1995).
 RL -1- CATALYTIC ACTIVITY: O-SUCCINYL-L-HOMOSERINE + L-CYSTEINE =
 CYSTATHIONINE + SUCCINATE (CAN ALSO USE HYDROGEN SULFIDE AND
 METHANETHIOL AS SUBSTRATES).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: Methionine biosynthesis; second step.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE TRANS-SULFURATION ENZYMES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: U32694; AAC21764.1; ALT_INIT.

DR HSSP: P00935; 1CS1.
 DR TIGR: H10086;
 DR InterPro: IPR000277; Cys_Met_MetA_PP.
 DR Pfam: PF01053; Cys_Met_MetA_PP; 1
 DR PROSITE: PS00868; Cys_Met_MetA_PP; 1
 DR Meteline biosynthesis; Lyase; Pyridoxal phosphate;
 KM Complete proteome.
 FT BINDING 200
 SQ SEQUENCE 369 AA; 40371 MW; D76A8A03FEAF7A14 CRC64;

Query Match 39.8%; Score 43; DB 1; Length 369;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 PDNVLISKDVGTAT 17
 DB 92 PDENVISDVYGGT 105

RESULT 13

AROL_YEAST STANDARD; PRT; 1588 AA.

AC P08566;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pentafunctional AROM polypeptide [includes: 3-dehydroquininate synthase
 (EC 4.2.3.4); 3-dehydroquininate dehydratase (EC 4.2.1.10) (3-
 dehydroquinase); Shikimate 5-dehydrogenase (EC 1.1.1.25); Shikimate
 kinase (EC 2.7.1.71); 3-phosphoshikimate 1-carboxyvinyltransferase
 (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP
 synthase) (EPSPS)].
 DE ARO1 OR AROM OR YDR127W OR YD9302.02.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NX NCBI_TaxID=4932;

RP SEQUENCE FROM N.A.
 RA MEDLINE=86076802; Pubmed=2825635;
 RA Duncan K., Edwards R.M., Coggin J.R.;
 RT "The pentafunctional aroM enzyme of *Saccharomyces cerevisiae* is a
 RT mosaic of monofunctional domains.";
 RL Biochem. J. 246:375-386(1987).

RC SEQUENCE FROM N.A.
 RA STRAIN=S288c / AB972;
 RA Oliver K., Harris D., Bartell B.G., Rajandream M.A.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 1-44 AND 1557-1588 FROM N.A.
 RA MEDLINE=89065111; Pubmed=2848727;
 RA Duncan K., Edwards R.M., Coggin J.R.;
 RT "The *Saccharomyces cerevisiae* Aro1 gene. An example of the
 RT co-ordinate regulation of five enzymes on a single biosynthetic
 RT pathway.";
 RL FEBS Lett. 241:83-88(1988).

CC -1- FUNCTION: THE AROM POLYPEPTIDE CATALYZES 5 CONSECUTIVE ENZYMATIC
 CC REACTIONS IN PRECHORISMATE POLYAROMATIC AMINO ACID BIOSYNTHESIS.
 CC CATALYTIC ACTIVITY: 3-deoxy-arabino-heptulosonate 7-phosphate = 3-
 CC dehydroquinolate + phosphate.
 CC -1- CATALYTIC ACTIVITY: 3-dehydroquinolate = 3-dehydroshikimate + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 5-dehydroshikimate +
 CC NADPH.

CC -1- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 CC PHOSPHATE + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -1- PATHWAY: SECOND TO SIXTH STEPS IN THE BIOSYNTHESIS OF CHORISMATE
 CC WITHIN THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE
 CC PATHWAY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC DEHYDROQUINATE SYNTHASE FAMILY.

CC -1- SIMILARITY: IN THE 3RD SECTION; BELONGS TO THE SHIKIMATE KINASE
 CC FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SHIKIMATE
 CC DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X06077; CAA29458.1; -
 CC EMBL: 248179; CAA88208.1; -
 CC EMBL: X13802; CAA32036.1; -
 CC EMBL: X13803; CAA32037.1; -
 CC PIR: A32519; BVBVAL.

DR HSSP: P07547; IDOS.
 DR SGD: S0002534; ARO1.
 DR InterPro: IPR002658; DHQ_synthase.
 DR InterPro: IPR001381; DHQ_synthase_I.
 DR InterPro: IPR001986; EPSP_synthase.
 DR InterPro: IPR000623; Shik_kinase.
 DR InterPro: IPR002907; Shikimate_DH.
 DR Pfam: PF00275; EPSP_synthase; 1.
 DR Pfam: PF01202; SKT; 1.
 DR Pfam: PF01487; DHQ_synthase_1; 1.
 DR Pfam: PF01488; Shikimate_DH; 1.
 DR Pfam: PF01761; DHQ_synthase; 1.
 DR PRINTS: PR01100; SHIKIMTKINASE.
 DR PRODOM: PD001867; EPSP_synthase; 1.
 DR TIGRFAMS: TIGR01093; aroD; 1.

DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
 DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
 DR PROSITE: PS01028; DEHYDROQUINATE_1; 1.
 DR PROSITE: PS01128; SHIKIMATE_KINASE; 1.
 DR K1 Aromatic amino acid biosynthesis; Multifunctional enzyme;
 DR Oxidoreductase; Lyase; Transferase; Kinase; NADP; ATP-binding.
 FT DOMAIN 1 392
 FT DOMAIN 404 866
 FT DOMAIN 887 1060
 FT DOMAIN 1061 1293
 FT DOMAIN 1306 1388
 FT ACT_SITE 853 853
 FT ACT_SITE 895 902
 FT NP_BIND 1198 1198
 FT ACT_SITE 1227 1227

FT ACT_SITE 1227 1227
 SQ SEQUENCE 1588 AA; 174754 MW; 9880EE2423CDE410 CRC64;
 Query Match 39.8%; Score 43; DB 1; Length 1588;
 Best Local Similarity 43.8%; Pred. No. 1e+02;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 LNDPNVLISKDVGTAT 17
 DB 645 INSEVIEEDASAT 660

RESULT 14
 ID POLN_SINDV STANDARD; PRT; 2512 AA.
 AC P03317;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nonstructural polyprotein (p270) [contains: Nonstructural protein
 DE NSP1; Nonstructural protein NSP2; Nonstructural NSP3; Nonstructural NSP4].
 OS Sindbis virus (strain HRSP).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 NX NCBI_TaxID=11034;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-84148439; PubMed-6322438;
RA Strauss E.G., Rice C.M., Strauss J.H.;
RT "Complete nucleotide sequence of the genomic RNA of Sindbis virus.";
RL Virology 133:92-110(1984).
RN [2]
RP SEQUENCE OF 1-54 FROM N.A.
RX MEDLINE-83268700; PubMed-6308269;
RA Ou J.H., Strauss E.G., Strauss J.H.;
RT "The 5'-terminal sequences of the genomic RNAs of several
RN alphaviruses.";
RL J. Mol. Biol. 168:1-15(1983).
RN [3]
RP SEQUENCE OF 1429-2512 FROM N.A.
RX MEDLINE-83299855; PubMed-6577423;
RA Strauss E.G., Rice C.M., Strauss J.H.;
RT "Sequence coding for the alphavirus nonstructural proteins is
RT interrupted by an opal termination codon.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:5271-5275(1983).
RN [4]
RP SEQUENCE OF 2431-2512 FROM N.A.
RX MEDLINE-83039346; PubMed-6291034;
RA Ou J.H., Rice C.M., Dalgaard L., Strauss E.G., Strauss J.H.;
RT "Sequence studies of several alphavirus genomic RNAs in the region
RT containing the start of the subgenomic RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:5235-5239(1982).
CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PFM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
CC BETWEEN THE CODONS FOR 1896-TYR AND 1897-LEU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02363; AAA96975.1; ALT_SEQ.
DR PIR: A03917; MNWVS.
DR MEROPS: C09.001;
DR InterPro: IPR002589; Alpp.
DR InterPro: IPR002620; Peptidase_C9.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
DR Pfam: PF01661; Alpp; 1.
DR Pfam: PF01707; Peptidase_C9; 1.
DR SMART: SM00506; Alpp; 1.
KW Polypeptide; Nonstructural protein; RNA-binding; Helicase.
FT CHAIN 1 540 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 1348 1896 NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 1897 2512 NONSTRUCTURAL PROTEIN NSP4.
FT CHAIN NONSTRUCTURAL PROTEIN NSP4.
SO SEQUENCE 2512 AA; 279546 MW; F3656FCBB8495726 CRC64;

Query Match 39.8%; Score 43; DB 1; Length 2512;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPDNVLSKDYVATKRVK 20
DB 877 NPCKNIEIDITGATKPK 894

```

```

DT 01-AUG-1992 (rel. 23, last sequence update)
DE 16-OCT-2001 (rel. 40, last annotation update)
DE Nonstructural polyprotein (p270) [Contains: Nonstructural protein
DE NSP1; Nonstructural NSP2; Nonstructural NSP3; Nonstructural NSP4].
OS Sindbis virus (subtype Ockelbo / strain Edsbyn 82-5).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=31699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91220725; PubMed-1673813;
RA Shirako Y., Niklasson B., Dalrymple J.M., Strauss E.G., Strauss J.H.;
RT "Structure of the Ockelbo virus genome and its relationship to other
RT Sindbis viruses.";
RL Virology 182:753-764(1991).
CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PFM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
CC BETWEEN THE CODONS FOR 1896-TYR AND 1897-LEU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M69205; AAA96972.1; ALT_SEQ.
DR PIR: A39991; MNW82.
DR MEROPS: C09.001;
DR InterPro: IPR002589; Alpp.
DR InterPro: IPR002620; Peptidase_C9.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR002588; V_methyltransf.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
DR Pfam: PF01660; Vmethyltransf; 1.
DR Pfam: PF01661; Alpp; 1.
DR Pfam: PF01707; Peptidase_C9; 1.
DR SMART: SM00506; Alpp; 1.
KW Polypeptide; Nonstructural protein; RNA-binding; Helicase.
FT CHAIN 1 540 NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 1348 1898 NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 1899 2514 NONSTRUCTURAL PROTEIN NSP4.
SO SEQUENCE 2514 AA; 279642 MW; 2F388CE32ACF5EDD CRC64;

Query Match 39.8%; Score 43; DB 1; Length 2514;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPDNVLSKDYVATKRVK 20
DB 877 NPCKNIEIDITGATKPK 894

```

Search completed: June 26, 2003, 00:50:34
Job time : 3.22353 secs

RESULT 15
POLN_SINDO STANDARD; PRT: 2514 AA.
AC P27283;
DT 01-AUG-1992 (rel. 23, Created)

RESULT 2

092NB5 PRELIMINARY: PRT: 1223 AA.
ID 092NB5
AC 092NB5
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-JUN-2002 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 130k-HMD (Fragment).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriia; CF8 group; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
CN 130k-HMD (Fragment).
RN 130k-HMD (Fragment).
RP SEQUENCE FROM N.A.
RC STRAIN-381;
RA MEDLINE-99143166; PubMed-9988746;
RA Shiba Y., Hayakawa M., Takiguchi H., Shirota T., Adiko Y.;
RT "Cloning, sequencing and characterization of the hemagglutinin-associated
RT short repeat found in Porphyromonas gingivalis multiple gene
RT products".
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL: AB019363; SAA34341.1.
DR InterPro: IPR000977; DNM_119988.
DR HSP: P5493; ICVR.
DR PIR: P01364; Peptidase_C25.
DR PIR: P01364; Peptidase_C25.
DR PROSITE: PS00697; DNM_LIGASE_A1; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 1223 AA: 131542 MW: 00232CD28A9F91B3 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1223;
Best Local Similarity 100.0%; Prd. No. 1 2e-08;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 ALNDPNTLSKDYTGATKRVY 21
DB 707 ALNDPNTLSKDYTGATKRVY 727

RESULT 3

096967 PRELIMINARY: PRT: 1358 AA.
ID 096967
AC 096967
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hemagglutinin.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriia; CF8 group; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
CN Hemagglutinin.
RN Hemagglutinin.
RP SEQUENCE FROM N.A.
RC STRAIN-381;
RA Han N., Leptine G., Wojciechowski L., Progniske-Fox A.;
RT "Cloning, sequencing and characterization of hagd, a member of the
RT HAGE multigene family in Porphyromonas gingivalis".
RT Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
RL EMBL: AB019363; SAA34341.1.
DR InterPro: IPR000977; DNM_119988.
DR HSP: P5493; ICVR.
DR PIR: P01364; Peptidase_C25.
DR PROSITE: PS00697; DNM_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1358 AA: 147102 MW: 47FC0B25506DEB8 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1358;
Best Local Similarity 100.0%; Prd. No. 1 3e-08;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 ALNDPNTLSKDYTGATKRVY 21
DB 850 ALNDPNTLSKDYTGATKRVY 870

RESULT 4

096967 PRELIMINARY: PRT: 1687 AA.
ID 096967
AC 096967
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-JUN-2002 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hemagglutinin/protease.
GN HAGE.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriia; CF8 group; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
CN Hemagglutinin/protease.
RN Hemagglutinin/protease.
RP SEQUENCE FROM N.A.
RC STRAIN-381;
RA Han N., Dong H., Progniske-Fox A.;
RT "Cloning and characterization of hagd from P. gingivalis 381".
RT Submitted (Sep-1999) to the EMBL/Genbank/DBJ databases.
RL EMBL: AB026946; AAD01810.1.
DR HSP: P5493; ICVR.
DR MEROPS: C25_001;
DR InterPro: IPR000977; DNM_119988.
DR HSP: P5493; ICVR.
DR PIR: P01364; Peptidase_C25.
DR PROSITE: PS00697; DNM_LIGASE_A1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1687 AA: 183702 MW: D0859516A39FE70 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1687;
Best Local Similarity 100.0%; Prd. No. 1 7e-08;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 ALNDPNTLSKDYTGATKRVY 21
DB 1179 ALNDPNTLSKDYTGATKRVY 1199

RESULT 5

051816 PRELIMINARY: PRT: 1704 AA.
ID 051816
AC 051816
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Arg-gingipain-1 proteinase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriia; CF8 group; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
CN Arg-gingipain-1 proteinase.
RN Arg-gingipain-1 proteinase.
RP SEQUENCE FROM N.A.
RC MEDLINE-9513080; PubMed-783031;
RA Travis J., Barr P.J., Fife K.N., Prochazka V., Klefer M.C.,
RT "Molecular cloning and structural characterization of the Arg-
RT gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
RT protease-inhibitor polypeptide".
RL EMBL: AB019363; SAA34341.1.
DR HSP: P5493; ICVR.
DR MEROPS: C25_001;
DR InterPro: IPR000977; DNM_119988.
DR PIR: P01364; Peptidase_C25.
DR PROSITE: PS00697; DNM_LIGASE_A1; UNKNOWN_1.
FT CHAIN 228
SQ SEQUENCE 1704 AA: 185436 MW: 6A34A013C2A676 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1704;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNDPNYLISKDYTGATKVKY 21
DB 1196 ALNDPNYLISKDYTGATKVKY 1216

RESULT 6

OS1839 PRELIMINARY; PRT; 1706 AA.

AC 051839; 051840;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
GN PRPRL.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirszenbaum L., Sotiropoulos C., Jackson C., Cleal S., Slakeski N., Reynolds E.C.;
RT "Complete nucleotide sequence of a gene prtr of Porphyromonas gingivalis W50 encoding a 132 kDa protein that contains an arginine-specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Biochem. Biophys. Res. Commun. 207:424-431(1995).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=96311339; PubMed=8713096;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes an arginine-specific thiol proteinase and multiple adhesins.";
RL Biochem. Biophys. Res. Commun. 224:605-610(1996).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Reynolds E.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Slakeski N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L26341; AAC18876.1; -.
DR HSSP: P95493; 1CVR.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM Protease; Signal.
FT SIGNAL. 1 23
FT CHAIN 228 1706 ARGININE-SPECIFIC THIOI. PROTEASE.
SQ SEQUENCE 1706 AA; 185626 MW; E8BD07C9813B844 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNDPNYLISKDYTGATKVKY 21
DB 1196 ALNDPNYLISKDYTGATKVKY 1216

RESULT 7
ID 051838 PRELIMINARY; PRT; 1706 AA.
AC 051838;

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Protease precursor.
GN PRPRL.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=96071894; PubMed=7591131;
RA Aduse-Opoku J., Milr J., Slaney J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease antigen (PRPRL) of Porphyromonas gingivalis W50.";
RL Infect. Immun. 63:4744-4754(1995).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Rangarajan M., Aduse-Opoku J., Slaney J.M., Young K.A., Curtis M.A.;
RT "The prprl and the prr2 arginine-specific protease genes of Porphyromonas gingivalis W50 produce five biochemically distinct enzymes.";
RL Mol. Microbiol. 23:0-0(1997).
DR EMBL: X82680; CA57997.1; -.
DR HSSP: P95493; 1CVR.
DR MEROPS: C25.001; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM Signal.
FT SIGNAL. 1 23
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E56DCD87DDA8CDD CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNDPNYLISKDYTGATKVKY 21
DB 1196 ALNDPNYLISKDYTGATKVKY 1216

RESULT 8

P72194 PRELIMINARY; PRT; 1723 AA.

AC P72194;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN SEQUENCE FROM N.A.
RC STRAIN=381;
RX MEDLINE=97044756; PubMed=8889827;
RA Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT "Cloning and sequencing of the gene encoding a novel lysine-specific cysteine proteinase (Lys-gingipain) in Porphyromonas gingivalis: structural relationship with the arginine-specific cysteine proteinase (Arg-gingipain).";
RL J. Biochem. 120:398-406(1996).
DR EMBL: D83258; BAA11870.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_ligase.

DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3. UNKNOWN_1.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1723 AA; 187261 MW; 5628963D231493EB CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ALNPDNYLSKDYTGATKVKY 21
 Db 1215 ALNPDNYLSKDYTGATKVKY 1235

RESULT 9
 ID P72197 PRELIMINARY; PRT; 1723 AA.
 AC P72197;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE lys-gingipain.
 GN KCP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
 RA Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
 RT "Molecular cloning and characterization of Porphyromonas gingivalis
 RT lys-gingipain";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL:U04691; AAA9810.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1723 AA; 186831 MW; 4508A7E50197CEBD CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLSKDYTGATKVKY 21
 Db 1215 ALNPDNYLSKDYTGATKVKY 1235

RESULT 10
 ID 052050 PRELIMINARY; PRT; 1732 AA.
 AC 052050;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE lysine specific cysteine protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W83;
 RX MEDLINE=98298016; PubMed=9632563;
 RA Lewis J.P., Macrina F.L.;
 RT "IS195, an insertion sequence-like element associated with protease
 RT genes in Porphyromonas gingivalis";
 RL Infect. Immun. 66:3035-3042(1998).
 DR EMBL: AF017059; AAC26523.1; -.
 DR MEROPS: C25.002; -.

DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3. UNKNOWN_1.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Protease.
 SO SEQUENCE 1732 AA; 187931 MW; B2337463D5CB5EAS CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ALNPDNYLSKDYTGATKVKY 21
 Db 1216 ALNPDNYLSKDYTGATKVKY 1236

RESULT 11
 ID 051817 PRELIMINARY; PRT; 1732 AA.
 AC 051817;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Porphyrypain.
 GN PRP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W12;
 RX MEDLINE=96213011; PubMed=8631659;
 RA Barkocy-Gallagher G.A., Han N., Patti J.M., Whitlock J.,
 RA Praguske-Fox A., Lantz M.S.;
 RT "Analysis of the prt gene encoding porphyrypain, a cysteine proteinase
 RT of Porphyromonas gingivalis";
 RL J. Bacteriol. 178:2734-2741(1996).
 DR EMBL: U04210; AAB06565.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1732 AA; 187875 MW; 654271DBE7BCAE4 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLSKDYTGATKVKY 21
 Db 1216 ALNPDNYLSKDYTGATKVKY 1236

RESULT 12
 ID 007442 PRELIMINARY; PRT; 1732 AA.
 AC 007442;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE lysine-specific cysteine proteinase.
 GN PRK.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W50;
 RX MEDLINE=99235907; PubMed=10219167;
 RA Stakeski N., Cleal S.M., Bhogal P.S., Reynolds E.C.;

* Characterization of a Porphyromonas gingivalis gene prtk that encodes a lysine-specific cysteine proteinase and three sequence-related adhesins."
 RT Oral Microbiol. Immunol. 14:92-97(1999).
 DR EMBL: U75366; AAB60809.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_Ligase_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA; 187914 MW; 45D5B9137391703 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVGTGATKVKY 21
 DB 1216 ALNPDNYLSKDVGTGATKVKY 1236

RESULT 13
 OY 09F4J0 PRELIMINARY; PRT; 925 AA.

AC 09F4J0: PRELIMINARY; PRT; 925 AA.
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Putative outer membrane protein PG57.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W50;
 RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,
 RT Hocking D., Webb E.;
 RT P. gingivalis polypeptides and nucleic acids."
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W50;
 RA Ross B.C., Czajkowski L., Hocking D., Margetts M., Webb E., Rothel L.,
 RT Patterson M., Agius C., Camuglia S., Reynolds E., Barr I.G.;
 RT "Identification of vaccine candidates from genomic analysis of
 RT Porphyromonas gingivalis."
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY007522; AAG24228.1; -.
 DR InterPro: IPR003961; FN_III.
 DR SMART: SM0060; FN3; 3.
 SQ SEQUENCE 925 AA; 103632 MW; 5FF2198D6914DAE8 CRC64;

Query Match 59.3%; Score 64; DB 2; Length 925;
 Best Local Similarity 60.0%; Pred. No. 0.12;
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 LNPDPNYLSKDVGTGATKVKY 21
 DB 367 VTPDPNYLTPKVEGAKRKVKY 386

RESULT 14
 OY 093922 PRELIMINARY; PRT; 654 AA.
 AC 093922: PRELIMINARY; PRT; 654 AA.
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE El-like protein.
 DE GSA7.
 GN Pichia pastoris (Yeast).
 OS Pichia pastoris: Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;

RP SEQUENCE FROM N.A.
 RX MEDLINE=99250146; PubMed=10233149;
 RA Yuan W., Stromhaug P.E., Dunn W.A. Jr.;
 RT "Glucose-induced autophagy of peroxisomes in Pichia pastoris requires
 RT a unique El-like protein."
 RL Mol. Biol. Cell 10:1353-1366(1999).
 DR EMBL: AF098976; AAD14610.1; -.
 DR InterPro: IPR000205; NAD_binding.
 DR InterPro: IPR000594; Thif_domain.
 DR Pfam: PF00899; Thif; 1.
 SQ SEQUENCE 654 AA; 74292 MW; 893937F5FB30D8F CRC64;

Query Match 52.8%; Score 57; DB 3; Length 654;
 Best Local Similarity 65.0%; Pred. No. 1.2;
 Matches 13; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

OY 1 ALNPDNYLSKDVGTGATKVKY 20
 DB 609 ALNENNYL-EDLTGLTKVK 626

RESULT 15
 OY 045985 PRELIMINARY; PRT; 334 AA.

AC 045985: PRELIMINARY; PRT; 334 AA.
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ZK1037.3 protein.
 GN ZK1037.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Basham V.M.;
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81142; CAB03502.1; -.
 DR InterPro: IPR002651; DUF32.
 DR Pfam: PF01748; DUF32; 1.
 SQ SEQUENCE 334 AA; 37719 MW; E8592D5A4D24FBA3 CRC64;

Query Match 50.9%; Score 55; DB 5; Length 334;
 Best Local Similarity 64.7%; Pred. No. 1.2;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LNPDPNYLSKDVGTGATKVKY 18
 DB 22 LNPDPVCSDDVTSHTK 38

Search completed: June 26, 2003, 00:53:27
 Job time : 7.42353 secs

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 15 AA:

Query Match 100.0%; Score 78; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDYTG 15
 DB 1 ALNDPNYLISKDYTG 15

RESULT 2

ID AAB49218 standard; peptide: 21 AA.

AC AAB49218;

DT 13-MAR-2001 (first entry)

DE Peptide #3.

KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 infection.

OS Unidentified.

PN WO200072875-A1.

XX 07-DEC-2000.

PF 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNITV SYDNEY.

PA Collyer CA, Hunter N, De Carlo AA;

PI Collyer CA, Hunter N, De Carlo AA;

DR MPI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist -
 PS Claim 11; Page 40; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 21 AA:

Query Match 100.0%; Score 78; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDYTG 15
 DB 1 ALNDPNYLISKDYTG 15

RESULT 3

ID AAB49217 standard; protein: 134 AA.

AC AAB49217;

DT 13-MAR-2001 (first entry)

DE Peptide used in the invention.

KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 infection.

OS Unidentified.

PN WO200072875-A1.

XX 07-DEC-2000.

PF 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNITV SYDNEY.

PA Collyer CA, Hunter N, De Carlo AA;

PI Collyer CA, Hunter N, De Carlo AA;

DR MPI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist -
 PS Claim 9; Page 98-99; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 134 AA:

Query Match 100.0%; Score 78; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDYTG 15
 DB 60 ALNDPNYLISKDYTG 74

RESULT 4

ID AAR96024 standard; Protein: 439 AA.

AC AAR96024;

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin hAgA HAREP4 product.

XX Haemagglutinin; hAgA; periodontal disease; vaccine; antibody;

KM HAREP4.
 OS Porphyromonas gingivalis strain 381.
 XX MO9617936-A2.
 PN 13-JUN-1996.
 PD 11-DEC-1995; 95WO-US16108.
 PF 09-DEC-1994; 94US-0353485.
 PR (UABR-) UAB RES FOUND.
 XX (UYFL) UNIV FLORIDA.
 PA Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI WPI; 1996-287181/29.
 XX N-PSDB; AAT30648.
 DR Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 114-115; 153pp; English.
 XX HAREP4 (AAR96024) is the product of the HAREP4 repeat unit (AAT30648)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP4 and other
 CC hga repeat unit products (see also AAR96021-23) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC animals against periodontal disease. Expression in Salmonella
 CC cells allows prodn. of live vaccine. HAREP4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 XX
 SQ Sequence 439 AA:
 Query Match 100.0%; Score 78; DB 17; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 5
 ID AAM69492 standard; Protein: 439 AA.
 AC AAM69492;
 XX 22-DEC-1998 (first entry)
 DT Haemagglutinin protein hga, HAREP4.
 DE Haemagglutinin protein; periodontal disease; vaccine; hga.
 KM Haemagglutinin protein; periodontal disease; vaccine; hga.
 XX Porphyromonas gingivalis.
 OS US5824791-A.
 PN 20-OCT-1998.
 PD 11-DEC-1995; 95US-0570311.
 PF 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX (UABR-) UAB RES FOUND.
 PA

PA (UYFL) UNIV FLORIDA.
 XX Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI WPI; 1998-582627/49.
 DR N-PSDB; AAV58879.
 DR Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))
 PS Claim 1; Column 139-144; 101pp; English.
 XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hga haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX
 SQ Sequence 439 AA:
 Query Match 100.0%; Score 78; DB 19; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 6
 ID AAR96021 standard; Protein: 450 AA.
 AC AAR96021;
 XX 04-SEP-1996 (first entry)
 DT P. gingivalis haemagglutinin hga HAREP1 product.
 DE Haemagglutinin; hga; periodontal disease; vaccine; antibody;
 KM HAREP1.
 XX Porphyromonas gingivalis strain 381.
 OS MO9617936-A2.
 XX 13-JUN-1996.
 PD 11-DEC-1995; 95WO-US16108.
 PF 09-DEC-1994; 94US-0353485.
 PR (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI WPI; 1996-287181/29.
 XX N-PSDB; AAT30645.
 DR Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 103-104; 153pp; English.
 XX HAREP1 (AAR96021) is the product of the HAREP1 repeat unit (AAT30645)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP1 and other
 CC hga repeat unit products (see also AAR96022-24) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or

XX The present invention relates to prophylaxis or treatment of infection
CC by a microorganism in a biological environment comprising iron, heme
CC or porphyrin. The treatment involves administering to the environment,
CC an agent which antagonizes the interaction between a molecule derived
CC from the microorganism and having an HA2 domain, and an HA2-binding
CC motif on a porphyrin containing molecule present in the environment.
CC Useful in the manufacture of a medicament for the prophylaxis and
CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
CC resulting from Porphyromonas gingivalis infection or infection by a
CC related microorganism.

SQ Sequence 15 AA;

Query Match 100.0%; Score 78; DB 22; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.3e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNPDNYLISKDVTG 15
Db 1 ALNPDNYLISKDVTG 15

RESULT 2

AAB49218

ID AAB49218 standard; peptide: 21 AA.

XX AC AAB49218;

DT 13-MAR-2001 (first entry)

XX Peptide #3.

DE HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
KW infection.

XX Unidentified.

PN WO200072875-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-AU00599.

XX 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNIV SYDNEY.

PI Collyer CA, Hunter N, De Carlo AA;

XX WPI; 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
PT administering a HA-2 antagonist -

PS Claim 11; Page 40; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
CC by a microorganism in a biological environment comprising iron, heme
CC or porphyrin. The treatment involves administering to the environment,
CC an agent which antagonizes the interaction between a molecule derived
CC from the microorganism and having an HA2 domain, and an HA2-binding
CC motif on a porphyrin containing molecule present in the environment.
CC Useful in the manufacture of a medicament for the prophylaxis and
CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
CC resulting from Porphyromonas gingivalis infection or infection by a
CC related microorganism.

SQ Sequence 21 AA;

Query Match 100.0%; Score 78; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 2e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNPDNYLISKDVTG 15
Db 1 ALNPDNYLISKDVTG 15

RESULT 3

AAB49217

ID AAB49217 standard; protein; 134 AA.

XX AC AAB49217;

XX 13-MAR-2001 (first entry)

XX Peptide used in the invention.

XX HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
KW infection.

XX Unidentified.

PN WO200072875-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-AU00599.

XX 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNIV SYDNEY.

PI Collyer CA, Hunter N, De Carlo AA;

XX WPI; 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
PT administering a HA-2 antagonist -

PS Claim 9; Page 98-99; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
CC by a microorganism in a biological environment comprising iron, heme
CC or porphyrin. The treatment involves administering to the environment,
CC an agent which antagonizes the interaction between a molecule derived
CC from the microorganism and having an HA2 domain, and an HA2-binding
CC motif on a porphyrin containing molecule present in the environment.
CC Useful in the manufacture of a medicament for the prophylaxis and
CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
CC resulting from Porphyromonas gingivalis infection or infection by a
CC related microorganism.

SQ Sequence 134 AA;

Query Match 100.0%; Score 78; DB 22; Length 134;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNPDNYLISKDVTG 15
Db 60 ALNPDNYLISKDVTG 74

RESULT 4

AAR96024

ID AAR96024 standard; Protein: 439 AA.

XX AC AAR96024;

XX 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hagA HAREP4 product.

XX Haemagglutinin; hagA; periodontal disease; vaccine; antibody;

KW Harep4.
 XX Porphyromonas gingivalis strain 381.
 OS
 XX WO9617936-A2.
 PN
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995; 95WO-US16108.
 PF
 XX 09-DEC-1994; 94US-0353485.
 PR
 XX {UABR-} UAB RES FOUND.
 PA {UYFL } UNIV FLORIDA.
 XX
 XX Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 XX WPI; 1996-287181/29.
 DR N-PSDB; AAT30648.
 DR
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 114-115; 153pp; English.
 XX
 XX Harep4 (AAR96024) is the product of the Harep4 repeat unit (AAT30648)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hgaA (see also AAR96030). Harep4 and other
 CC hgaA repeat unit products (see also AAR96021-23) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC animals against periodontal disease. Expression in Salmonella
 CC cells allows prodn. of live vaccine. Harep1-4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 XX
 XX Sequence 439 AA;
 SQ
 Query Match 100.0%; Score 78; DB 17; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLSKDVGTG 15
 DB 251 ALNPDNYLSKDVGTG 265
 RESULT 5
 AAW69492
 ID AAW69492 standard; Protein; 439 AA.
 AC
 XX AAW69492;
 XX
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein hgaA, Harep4.
 XX
 XX Haemagglutinin protein; periodontal disease; vaccine; hgaA.
 KW
 XX Porphyromonas gingivalis.
 OS
 XX US5824791-A.
 PN
 XX 20-OCT-1998.
 PD
 XX 11-DEC-1995; 95US-0570311.
 PF
 XX 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 XX {UABR-} UAB RES FOUND.
 PA

PA {UYFL } UNIV FLORIDA.
 XX
 XX Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 XX WPI; 1998-582627/49.
 DR N-PSDB; AAV58879.
 XX
 XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s)
 PS Claim 1; Column 139-144; 101pp; English.
 XX
 XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hgaA haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX
 XX Sequence 439 AA;
 SQ
 Query Match 100.0%; Score 78; DB 19; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLSKDVGTG 15
 DB 251 ALNPDNYLSKDVGTG 265
 RESULT 6
 AAR96021
 ID AAR96021 standard; Protein; 450 AA.
 XX
 XX AAR96021;
 XX
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haemagglutinin hgaA Harep1 product.
 XX
 XX Haemagglutinin; hgaA; periodontal disease; vaccine; antibody;
 KW Harep1.
 XX
 XX Porphyromonas gingivalis strain 381.
 OS
 XX WO9617936-A2.
 PN
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995; 95WO-US16108.
 PF
 XX 09-DEC-1994; 94US-0353485.
 PR
 XX {UABR-} UAB RES FOUND.
 PA {UYFL } UNIV FLORIDA.
 XX
 XX Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 XX WPI; 1996-287181/29.
 DR N-PSDB; AAT30645.
 DR
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 103-104; 153pp; English.
 XX
 XX Harep1 (AAR96021) is the product of the Harep1 repeat unit (AAT30645)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hgaA (see also AAR96030). Harep1 and other
 CC hgaA repeat unit products (see also AAR96022-24) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC

CC incorporating AG-1.
 XX Sequence 1704 AA;
 Alignment Scores:
 Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservatives: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 16 Gaps: 0

US-09-980-370-5 (1-402) x AAR70188 (1-1704)

QY 1 GCAGACTTCACGGAACGTTCCGAGTCTTCTACTCATGGAGAGCCAGCGGAGTGAAGT 60
 DB 1137 AlaaspPheThrGluThrPheGluSerThrHisGlyGluAlaProAlaGluTrpThr 1156
 QY 61 ACTATCGATGCCGATGGCGATGGTGAAGGTTGGCTCTGTCTCTCCGACCAATTCGAC 120
 DB 1157 ThrileAspAlaaspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1176
 QY 121 TGGCTCACAGCTCATGGCGGACCAAGCTAGTAAGCTCTTCTCATGGAATGGAATGCT 180
 DB 1177 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1196
 QY 181 TTGAATCCTGATAACTATCTCTCAAGGATGTTACAGCGCCAGCAAGGTAAGTAC 240
 DB 1197 LeuAsnProAspAsnTyrLeuileSerLysAspValThrGlyAlaThrLysValLysTyr 1216
 QY 241 TACTATCCAGTCAACGACGGTTCCTCCGCGGATCACTATGCGGTGATGATCTCCAGACG 300
 DB 1217 TyrTyrAlaValAsnAspGlyPheProGlyAspHisTyrAlaValMetileSerLysThr 1236
 QY 301 GCGCAGCAAGCTTCGGTCTTCCAGGAGCAATGGCGCC 402
 DB 1237 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1256

RESULT 5
 AAW34843
 ID AAW34843 standard; protein; 1704 AA.
 AC AAW34843;
 DT 03-JUN-1998 (first entry)
 DE Arg-gingipain high molecular weight prepolyprotein sequence.
 KW Arg-specific gingipain protease; gingivalis; periodontal disease;
 OS Porphyromonas gingivalis.
 FH Key Location/Qualifiers
 FT Protein 1..227
 FT /note= "precursor protein"
 PN W09734629-1.
 XX 25-SEP-1997.
 XX -21-MAR-1997; 97WO-US04635.
 XX 22-MAR-1996; 96US-0013945.
 XX (MORE-) MOREHOUSE SCHOOL MEDICINE.
 XX (UYGE-) UNIV GEORGIA RES FOUND INC.
 PI Genco CA, Potempa J, Travis J, Genco C;

XX WPI; 1997-479993/44.
 DR N-PSDB; AAT93872.
 XX Porphyromonas gingivalis Arg-specific gingipain protease peptide(s)
 PT - useful for protecting animals and humans from gingivalis and
 PT periodontal diseases
 XX Disclosure; Pages 68-73; 95pp; English.
 XX The present sequence represents an arginine-specific protease of
 CC Porphyromonas gingivalis. The following peptides, derived from Arg-
 CC and Lys-specific high molecular weight proteases, offer protection
 CC against infection: YTYTVYRDGK IKEGLTATTE DDGVTATGNHE YCVEKYTAGS VSPKVC
 CC (I); YTPVEEKONG RMIVIVAKKY (II); QLPFIFDVAC VNGDFLFSMP CFAEALMRAO
 CC (III); GEPNPQVPS NLTTATGOGK VTLKWDAPSTK (IV); GNHEYCVVEVK YTAGCVSPKVC
 CC KDVTY (V); RMFMNYEPGR YTPVEEKONG (VI); TTAGFEDTYK RMFMNYEPGR (VII);
 CC DYTATVYRDG TTIKEGLTAT TFEEDGVATG NMEYCVVCWKY TAGVSPKVC (VIII); YTYTVYRDGT
 CC KIKEGLTATTF EEDG (IX); ROGTKIKEGL TATTFEEDGV ATGN (X); KIKEGLTATT
 CC FEEDGVATGN HEY (XI); KWDAPNGTNP PNPN PNPNP PGTTLTSE (XII); and
 CC YTPVEEKONG RMIVIVAKKY (XIII). They are used in vaccines to protect
 CC animals, including humans, from gingivitis and/or periodontal
 CC diseases.
 XX Sequence 1704 AA;
 Alignment Scores:
 Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservatives: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 16 Gaps: 0

US-09-980-370-5 (1-402) x AAW34843 (1-1704)

QY 1 GCAGACTTCACGGAACGTTCCGAGTCTTCTACTCATGGAGAGCCAGCGGAGTGAAGT 60
 DB 1137 AlaaspPheThrGluThrPheGluSerThrHisGlyGluAlaProAlaGluTrpThr 1156
 QY 61 ACTATCGATGCCGATGGCGATGGTGAAGGTTGGCTCTGTCTCTCCGACCAATTCGAC 120
 DB 1157 ThrileAspAlaaspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1176
 QY 121 TGGCTCACAGCTCATGGCGGACCAAGCTAGTAAGCTCTTCTCATGGAATGGAATGCT 180
 DB 1177 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1196
 QY 181 TTGAATCCTGATAACTATCTCTCAAGGATGTTACAGCGCCAGCAAGGTAAGTAC 240
 DB 1197 LeuAsnProAspAsnTyrLeuileSerLysAspValThrGlyAlaThrLysValLysTyr 1216
 QY 241 TACTATCCAGTCAACGACGGTTCCTCCGCGGATCACTATGCGGTGATGATCTCCAGACG 300
 DB 1217 TyrTyrAlaValAsnAspGlyPheProGlyAspHisTyrAlaValMetileSerLysThr 1236
 QY 301 GCGCAGCAAGCTTCGGTCTTCCAGGAGCAATGGCGCC 402
 DB 1237 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1256

RESULT 6
 AAW67396
 ID AAW67396 standard; Protein; 1704 AA.
 XX AAW67396;
 AC AAW67396;
 XX 25-APR-2000 (first entry)
 XX Arg-gingipain-2 amino acid sequence.

